

**Universidad Central de Venezuela**  
**Facultad de Humanidades y Educación**  
**Escuela de Artes**  
**Mención Cine**

**La animación a través de las tecnologías de los videojuegos:  
Uso de la captura de movimiento sin marcadores.**  
Trabajo Especial de Grado en Teoría para optar a la Licenciatura en Artes

Tutora: Grecia Almeida  
Oscar A. Febres Cordero G.  
C.I. 19.868.233

Caracas, julio de 2016



## **Agradecimiento**

A mi madre y a mi Abuela, las dos Mercedes más importantes en mi vida, sin su guía jamás hubiera encontrado el camino para alcanzar mis metas.

A mi padre Steven, que sin su apoyo y sacrificios no habría logrado mucho en la vida.

Fabiana Aparicio agradezco que siempre hayas estado a mi lado en este proyecto.

Müller, que sin importar que tan estresado estaba, siempre me sacabas una sonrisa.

A cada uno de los profesores de la Escuela de Arte de la UCV, quienes a pesar de la situación que vivimos siguen apostando por nuestra formación.

A mi tutora Grecia por haber creído en este proyecto y aceptarme como su tesista.

A todos mis amigos y familiares que me han apoyado en este trayecto.



## Contenido

Agradecimiento .....	i
Resumen .....	vii
Introducción.....	1
Capítulo I: La Captura de movimiento en la animación.....	5
1.1.- Inicios de la animación.....	5
1.2.-Predecesores de la captura de movimiento y los principios de la animación. ....	8
1.3.- Avances tecnológicos en la animación.....	16
1.4.- Origen de la captura de movimiento: .....	24
1.5.- El proceso de captura de movimiento: .....	31
1.5. 1.- Tipos de Sistemas de captura de movimiento .....	32
Capítulo II: Las tecnologías de los videojuegos como Sistema óptico de Captura de movimientos. ....	53
2.1.- Accesorios de videojuegos, como sistemas de captura de movimiento.....	54
2.1.1. El Playstation Eye y Move. ....	54
2.1.2.- El Kinect para la captura de movimiento.....	56
2.1.3.- Múltiples cámaras (RGB) o múltiples Sensores de Profundidad vs (RGB-D) ¿cuál es el más viable para la captura?.....	62
2.2. Software con que funciona el Kinect.....	64
2.2.1.- iPiSoft Captura de Movimiento para las Masas. ....	64
2.2.2. –Brekel3D.....	72
Capítulo III: La producción en la captura de Movimiento. ....	75
3.1.- Preproducción. ....	75
3.1.1. – Elaboración del Guión .....	76
3.1.2. – Elaboración del Guion Grafico.....	76

3.1.3. – Decidir si la captura de movimiento es adecuada.....	77
3.1.4. – Lista de tomas.....	81
3.1.5. – Motion Capture Blueprints o Planos de Captura de Movimiento .....	82
3.1.6. – Realización del Animatic.....	92
3.1.7. - El talento y audiciones .....	93
3.1.8. – Itinerario de captura:.....	94
3.2. – Trabajo de pre-captura de movimiento.....	95
3.2.1. – Preparar el área para la captura de movimiento: .....	95
3.2.2. – Ensayos.....	95
3.2.3. – Designar y preparar la utilería. ....	96
3.3. – Captura de Movimiento .....	96
3.3.1. – Calibrar el equipo .....	97
3.3.2. – Toma de prueba .....	100
3.3.3. – Tener una referencia tanto de audio como de video.....	100
3.3.5. – Organización de los archivos grabados. ....	101
3.4. – Post-captura de movimiento .....	101
3.4.1. – Revisar la data y mejorarla .....	101
3.4.2. – Aplicar la data al personaje 3D.....	104
3.5. –Post-producción y Rendering. ....	104
Capitulo IV: Captura de Movimiento en la práctica.....	105
4.1.- Fase de preproducción.....	105
4.1.1.- Elaboración del guion .....	105
4.1.2.- Pruebas de captura de movimiento.....	106
4.1.3.- El plano de Captura de movimiento.....	114
4.2.- Fase de producción.....	118

4.2.1 Producción Tradicional.....	118
4.2.2 Producción de Captura de Movimiento.....	119
4.2.3 Proceso de Animación.....	120
4.3.- Fase de Post-producción.....	122
Conclusiones.....	125
Glosario .....	133
Bibliografía.....	137
Anexos.....	145





## **Resumen**

El presente trabajo centro la atención en el uso de nuevas tecnologías, cuyos aportes han atenuado el esfuerzo en tiempo y recursos invertidos en la animación para la producción cinematográfica, tuvo como propósito determinar la viabilidad del sistema de captura de movimiento sin marcadores, basado en tecnología de videojuegos, específicamente en el uso de XBOX Kinect para la realización de productos cinematográficos animados.

Este es un estudio enmarcado en la investigación documental y de carácter exploratorio, debido a que buscó ubicar y delimitar los avances tecnológicos que se han logrado en la animación, su evolución en el tiempo, el proceso de captura de movimiento, los tipos de sistemas existentes para dicha captura, pero su énfasis estuvo en las tecnologías de videojuegos, empleando el Xbox Kinect y el software iPiSoft en la animación. Se realizó una prueba para comprobar las bondades y limitaciones de esta tecnología en la producción de animación. Por ende puede afirmarse que, el propósito del estudio fue logrado, se comprobó la viabilidad del sistema de captura de movimiento sin marcadores, basado en tecnología de videojuegos.

**Palabras claves:** animación, captura de movimiento, tecnologías de videojuegos, Xbox Kinect.



## Introducción

La animación es la forma más antigua de realizar películas. Consiste en crear la ilusión del movimiento con una serie de dibujos e imágenes. En la animación se construye un movimiento inexistente en la realidad, se considera como una ilusión óptica que dota de vida a un objeto. Existen diferentes técnicas para realizar animación, desde los dibujos animados, la fotografía, hasta la incorporación de los objetos de la realidad. Hoy con el desarrollo de la tecnología de la computación y digitalización, se ha facilitado la producción de películas, tanto animadas como de actores reales; sin embargo, la animación sigue requiriendo mucho tiempo, esfuerzo humano y recursos, por ello es necesario conocer qué se está realizando en esta área de la cinematografía, qué tecnologías existen para agilizar la producción y atenuar la inversión de tiempo y esfuerzo. Frente a este planteamiento surgió la inquietud por realizar el presente trabajo, el cual buscó investigar sobre el uso de nuevas tecnologías, se centró específicamente en el proceso de captura de movimiento, usado en los videos juegos, debido a que esta tecnología puede simplificar la producción de cortos con animación.

A fin de concretar esta intención fue necesario ejecutar las siguientes acciones: Primero, se realizó una investigación sobre la historia de la animación, así como también, sobre los avances tecnológicos que se han logrado en su producción, en especial lo relacionado con el proceso de captura de movimiento, se indagó sobre cómo estos avances han cambiado la forma y el contenido de hacer cine. Segundo, se explicó el Xbox Kinect como sistema óptico de captura de movimiento y se describieron los proyectos que se han realizado, utilizando esta técnica de captura de movimiento, enfatizándose como esto sirve como un apropiado indicador de las tendencias mundiales en este campo. Tercero, se explicó el proceso de captura de movimiento a través de las tecnologías de los videojuegos. Cuarto, se realizaron ensayos de captura de movimiento, a fin de tener una muestra concreta de su uso, y así poder mostrar las ventajas y desventajas de esta tecnología en la producción de cortos.

Acorde a los propósitos que se persiguieron en este estudio, se podría decir que el mismo se enmarca entre los estudios de investigación documental y de tipo exploratorio, debido a que se ubicaron y delimitaron cuáles son los avances tecnológico que se han logrado en la animación, su evolución en el tiempo, haciendo especial hincapié en el proceso de captura de movimiento,

lo cual se hizo mediante la consulta de fuentes bibliográficas en físico y digital. Pero además, se probó cómo el proceso de captura de movimiento a través de las tecnologías de los videojuegos se puede utilizar en la producción de cortos, lo cual permitió explorar y verificar sus bondades y limitaciones.

Las interrogantes que impulsaron el interés de la presente investigación fueron las siguientes: ¿Cuáles avances tecnológicos existen en la animación? ¿Cuáles son los antecedentes tecnológicos de la captura de movimiento en la animación? ¿Cómo funciona la Captura de Movimiento? ¿Cómo se hace la Captura de Movimiento? ¿Con el uso de esta tecnología se simplifica la animación en la producción de cortos? ¿Estos métodos son realmente más económicos y accesibles para fomentar la animación en la producción de cortos? La realización de la presente investigación permitió dar respuesta a las anteriores interrogantes.

Las razones que justifican y muestran la importancia de esta investigación son varias, entre ellas están las siguientes: lo primero es que en Venezuela hay escasa bibliografía sobre el tema, tampoco pareciera que la hay en abundancia en el idioma español, por lo tanto este estudio es un aporte no solo por el contenido que se desarrolló, sino por las fuentes bibliográficas que se consultaron y que aquí se dejaron referenciadas. En segundo lugar, al probar las nuevas tecnologías para la animación, se está contribuyendo a difundir los aspectos teóricos y prácticos del empleo de la captura de movimiento, a través de las tecnologías de los videojuegos y se mostró la factibilidad del uso de las mismas en la producción de cortos.

El trabajo que aquí se presenta, se estructuró en estos cuatro capítulos: en el primero se aborda cinco gruesos tópicos que pasean al lector por la captura de movimiento en la animación, por ende se alude a los inicios de la animación, destacándose los aportes de diverso estudiosos del campo, así mismo se describen aspectos esenciales relativos a los predecesores de la captura de movimiento y los principios de la animación, explicándolos con ejemplos ilustrativos para su aplicación, Se esbozan los principales avances tecnológicos en la animación, con lo cual se pudo precisar la necesaria simbiosis cine y tecnología, mostrándose e ilustrando los procesos que dieron origen a la captura de movimiento, se cierra este capítulo con el proceso de captura de movimiento, allí se describen los tipos de Sistemas de captura de movimiento. El capítulo dos es menos prolifero que el anterior, su foco de atención son Las tecnologías de los

videojuegos como Sistema óptico de Captura de movimientos, se describen y explican los diversos accesorios de videojuegos, incluyendo iPiSoft Captura de Movimiento para las Masas y el Brekel3D. El capítulo tres está dedicado a los procesos propiamente dichos de la producción en la captura de movimiento y el último capítulo se destinó a mostrar en la práctica cómo se efectúa la captura de movimiento, se revisan y ejemplifican las tres fases básicas de toda producción: la de preproducción, producción y postproducción con sus respectivos pasos a cumplir. De ahí que, se asevere que lo propuesto en el proyecto al iniciar el presente trabajo fue realmente logrado.



## Capítulo I: La Captura de movimiento en la animación.

El tema central de este trabajo es el uso de la captura de movimiento en la animación a través de las tecnologías de los videojuegos, para poder hablar adecuadamente sobre ello, se debe comenzar primero por conocer un poco sobre la historia del cine y sus predecesores, específicamente sobre los primeros intentos de animación.

### 1.1.- Inicios de la animación

La animación tuvo sus inicios mucho antes de las caricaturas conocidas hoy en día, si nos vamos al origen etimológico de la palabra animar, según la Real Academia Española, es “tr. Dar vida o animación a una obra de arte.”, esto es lo que fue exactamente la animación antes del surgimiento del cine. Animar era crear esa ilusión que le daba movimiento a objetos que por su propia cuenta no lo tenían. El hombre desde sus inicios ha estado fascinado por el movimiento, así por ejemplo, ha sentido gran atracción y curiosidad por el paso del día a la noche, las estrellas, el poder volar, dar vida a una imagen; al respecto no es difícil percibir a los primeros humanos bajo la luz del fuego danzante, cuando imaginaban los dibujos que realizaron en las paredes de las cavernas, moviéndose y cobrando vida (Azéma & Rivère, 2012). Esto es lo que se conoce por animación, es el proceso de dar la sensación de movimiento a imágenes u objetos inanimados. La animación se considera normalmente una ilusión óptica, y Romaguerra i Ramio indican que:

Es claramente una técnica, una manera genuina de crear movimiento mediante unas imágenes realizadas con una técnica, tan específica como fotografiar uno a uno, consecutivamente...que permite narrar historias...

(Ramió, 1999, p. 45)

Es oportuno aclarar que la animación es más antigua de lo que muchos se imaginan, es anterior al cine, según el artículo publicado en el CAIS, el CHTHO (2008) de Irán, el primer artefacto para la animación fue encontrado en restos arqueológicos en Shahr-e Sukhteh, Irán, en donde se encontró un tazón con 5 imágenes de una cabra saltando para comer de un arbusto. El tazón parece encajar en algún artefacto que le facilitaba girar, haciendo un efecto similar al de

un phenakistoscopio y un Zootropo, a pesar de ser creados como juguetes en la época victoriana funcionaban con los mismos principios que permitieron conocer el cine.

Durante ese periodo se realizaron diferentes tipos de juguetes y atracciones, los cuales tenían animaciones rudimentarias, algunas de ellas perduraron hasta los inicios del cine. La más importante de estas atracciones, por lo que se convirtieron en las películas animadas, fue la derivación del praxinoscopio, que fue inventado por el francés Charles-Émile Reynaud en 1877. Este aparato es similar al zootropo, el cual utilizaba una tira de imágenes colocadas alrededor de la superficie interior de un cilindro que gira. Pero el praxinoscopio se diferencia de éste, porque reemplaza las ranuras estrechas del zootropo con un círculo interno de espejos, colocados de forma que los reflejos de las imágenes parezcan más o menos estacionarios en relación con la rueda que gira. Por consiguiente, alguien que observa los espejos ve una rápida sucesión de imágenes que producen la ilusión de movimiento, pero con una imagen más brillante y menos distorsionada que la mostrada con el zootropo (Auzel & Tchernia, 2000).

Reynaud en el año 1888 patenta el *Théâtre Optique* o teatro óptico, basado en el praxinoscopio, el cual permitía ofrecer un espectáculo a partir de la proyección de dibujos animados móviles, dispuestos en una banda flexible provista de perforaciones, las cuales discurren con una frecuencia de 15 imágenes por segundo frente a un sistema de lentes e iluminación, basado en la linterna mágica. Y el 28 de octubre 1892, se dio la primera presentación pública de un espectáculo de imágenes en movimiento en el Museo Grévin de París. El espectáculo, considerado como *Pantomimes Lumineuses*, incluyó tres caricaturas, *Pauvre Pierrot* (1892), *Un bon bock* (1892), y *Le Clown et ses chiens* (1892), cada una compuesta de 500 a 600 imágenes, pintadas individualmente y con una duración de unos 15 minutos. Se puede decir que, Reynaud sembró la semilla de la historia de los dibujos animados(Auzel & Tchernia, 2000).

El historiador señala que, el 28 de diciembre de 1895 los hermanos Lumière realizan la primera exhibición cinematográfica en el Salon Indien du Grand Café de Paris, allí se presentaron una serie de 10 cortos de una duración menor a un minuto (Nowell-Smith, 1999). Siendo el corto más popular: la reconocida película del tren llegando a la estación, with *L'Arrivée d'un Train en Gare de la Ciotat*, lo cual literalmente significa la llegada de un tren a



la Ciotat, pero es conocida comúnmente como La llegada del tren a la estación, esto causó toda una conmoción en el café, al punto que algunos espectadores saltaron de sus asientos para esquivar al tren. Este espectáculo terminó eclipsando al teatro óptico de Reynaud. Pero el mismo se continuó presentando en el Museo Grévin hasta 1900 y para el momento de su cierre más de 500.000 personas lo habían visto. A partir de este momento, la idea de animar dibujos tomó un segundo plano, ya que el mundo se vio fascinado por la idea de ver la realidad a través de las pantallas de cine (Nowell-Smith, 1999). Y pasaron casi 10 años para que los dibujos animados fuesen proyectados a través del film.

Tres años después del debut del aparato de los Lumière, en Estados Unidos surgen compañías productoras que utilizan la tecnología de cámaras y proyectores desarrollados por Thomas A. Edison. Una de estas es Vitagraph Studios fundada en 1897 por J. Stuart Albert E. Smith, y la cual llegará a ser una de las compañías más prolíferas en los inicios del cine. Para 1898 producen los que varios expertos consideran que es la primera película animada en la historia, con *The Humpty Dumpty Circus* en ella utilizan la técnica que en la actualidad se denomina Stop-Motion, la cual consiste en aparentar el movimiento de objetos estáticos por medio de una serie de imágenes fijas sucesivas, esto permite crear una animación tridimensional, la cual se utilizó a lo largo del siglo XX para la creación de efectos especiales, o películas animadas en su totalidad.

Pierre Courtet-Cohl escribe sobre la vida y obra de Emile Cohl, quien fue un caricaturista francés, contratado por una productora francesa para producir cortos graciosos, con actores (Courtet-Cohl, Génin, & Takahata, 2008). En 1907 había comenzado a surgir distintos tipos de animaciones, principalmente la animación de objetos inanimados, a través de la animación cuadro por cuadro; estos se volvieron populares, y por esta razón el estudio le encarga a Emile Cohl que realice algo similar, a objeto de poder sacarle provecho a esta novedad. Cohl ve en su pasado como caricaturista una inspiración, es una animación sencilla, que simula como si fuera dibujado en un pizarrón con tiza. Cuando Cohl comenzó a dibujar el corto no tenía un guion, ya que quería que fuera inspirado por los mismos dibujos. Él utilizó una caja de luz para poder calcar el dibujo anterior y así poder realizar el movimiento sin que tuviera saltos; a medida que iba terminado el dibujo hacia una doble exposición del dibujo. Es decir que cada dibujo duraba 2 cuadros, luego lo utilizaba como base para realizar el próximo dibujo.

Este proceso culmina con una serie de 3 películas animadas, la primera de las cuales fue estrenada el 17 de agosto de 1908; la película es relativamente corta con una duración no mayor de un minuto y medio, pero estaba constituida por 700 imágenes. Fantasmagoría (Cohl, 1908) es considerada por los historiadores como la primera película completamente animada, empleando lo que hoy en día llamamos animación tradicional, que consiste en calcar y luego fotografiar los dibujos, a fin de evitar saltos en la imagen y movimientos bruscos.

La animación tradicional, a veces llamada animación dibujada a mano, fue el proceso más utilizado para hacer películas animadas durante el siglo XX. Se basa en los conocimientos adquiridos a finales del siglo XIX, por los predecesores del cine, de quienes se habló anteriormente. En esencia, se trata del fenómeno *phi*, es la ilusión óptica que ocurre debido a la persistencia retiniana que lleva a nuestro cerebro a percibir un movimiento continuo en donde hay una sucesión de imágenes. Para crear la ilusión de movimiento, cada dibujo es ligeramente diferente al anterior, cada fotograma individual de una película animada tradicional son fotografías, primero dibujados en papel, los dibujos de los animadores son trazadas o fotocopiados en hojas de acetato transparentes, llamadas celdas, que se rellenan con pinturas de colores o tonos asignados en el lado opuesto a los dibujos lineales. Las celdas de los personajes terminados son fotografiadas una por una contra un fondo pintado por una cámara, hasta terminar la película (Courtet-Cohl et al., 2008).

## **1.2.-Predecesores de la captura de movimiento y los principios de la animación.**

Al mismo tiempo que los avances en la animación y el cine florecían en Europa y Norteamérica, Eadweard Muybridge y Étienne-Jules Marey para 1860 iniciaron sus estudios del movimiento, a ellos se les puede considerar como los pioneros de la captura del movimiento, debido a que ellos establecieron las bases para su desarrollo futuro. Según Clegg escritor de la biografía del reconocido fotógrafo inglés Eadweard Muybridge, quien se hizo famoso por sus fotografías del Valle Yosemite en la década de los 1860, fue contratado por Leland Stanford para que comprobara una duda que él tenía y le permitiera ganar una apuesta que había hecho; Stanford quería específicamente que fotografiara a un caballo cuando cabalga, a fin de comprobar si en algún momento las cuatro patas del animal estaban en el aire al mismo tiempo (Clegg, 2007). Muybridge logró tomar una foto y comprobar que las cuatro patas del caballo estaban en el aire al mismo tiempo en un punto de su galope. A partir de esto, Muybridge

se ve intrigado por la idea de capturar el movimiento de un animal y sigue trabajando en la misma; quiso tener una serie de fotos que le permitieran ver todo el movimiento del animal. Para ello utilizó una serie de cámaras que se iban activando automáticamente con el galopar del caballo, esto lo cautivó tanto que lo llevó a desarrollar la técnica, hasta perfeccionarla.

Según Clegg, fue el 15 de junio de 1878, cuando logra realizar una serie de 24 fotos en las cuales se puede ver al caballo Sally Gardner en pleno cabalgar. Estas fotografías se proyectaron a la prensa a través de un aparato de su propia invención, el *Zoopraxiscopio*, situación que llevó a Muybridge a tener un reconocimiento en el mundo académico, además lograr un financiamiento por parte de la Universidad de Pennsylvania y publicar el libro *Animal Locomotion: An Electro-photographic Investigation of Consecutive Phases of Animal Movements, 1872–1885*. Esta investigación consistió en 781 placas de vidrio de fototipia, las placas de fototipia median 19 por 24 pulgadas cada una y se encuentran enmarcadas en 36 por 36 pulgadas, y cada una contenía la serie de movimientos capturados; el número total de imágenes fueron aproximadamente 20.000. Las placas publicadas incluyen 514 de hombres y mujeres en movimiento, 27 placas de movimiento anormal masculino y femenino, 16 de niños, 5 placas del movimiento de la mano de un adulto, y 219 con sujetos animales (Clegg, 2007). Este libro sirvió de referencia para artistas y académicos que deseaban estudiar el movimiento real, tanto de personas como animales.

Braun habla sobre el trabajo de Étienne-Jules Marey, quien fue un fisiólogo francés, que entre sus muchos intereses estaba la fotografía, por lo que desarrolla lo que él llama cronofotografía. Él comenzó a indagar en la fotografía del movimiento después de ver los trabajos de Muybridge. A diferencia de Muybridge, quien utilizaba múltiples cámaras, Marey pensaba que el uso de tantas cámaras era poco científico. Pero no fue sino hasta 1882 cuando logra desarrollar la pistola cronofotográfica, la cual utiliza una rueda para rápidamente tomar una secuencia de fotos; la referida pistola era capaz de tomar 12 cuadros por segundo, y todos los cuadros eran grabados en una misma imagen. Esto permite tomar medidas y estudiar mejor el movimiento. Publicó en 1894 el libro *Le Mouvement*, donde recopila toda la información que había desarrollado a lo largo de los años. Lo más interesante son sus experimentos con la cronofotografía geométrica, en las cuales Marey diseñó unos trajes especiales, eran negros y

tenían un material refractante, con su utilización y una iluminación especial se podía capturar el movimiento humano, pero sin apreciarse el cuerpo del mismo (Braun, 1995).

Es innegable la influencia que los trabajos de Marey tuvieron en esta área, fueron marcadamente influyentes en la captura de movimiento, en especial si se comparan con los trajes modernos, con los videos de la captura de movimiento modernos que utilizan trajes oscuros con marcadores refractante (Ver figura 1.1 ). Es por esta razón que, Kitagawa y Windsor (2008) consideran que tanto Muybridge como Marey son los precursores de la captura de movimiento; además que sus logros en el área de la fotografía sirvieron como bases del surgimiento de la captura de movimiento.

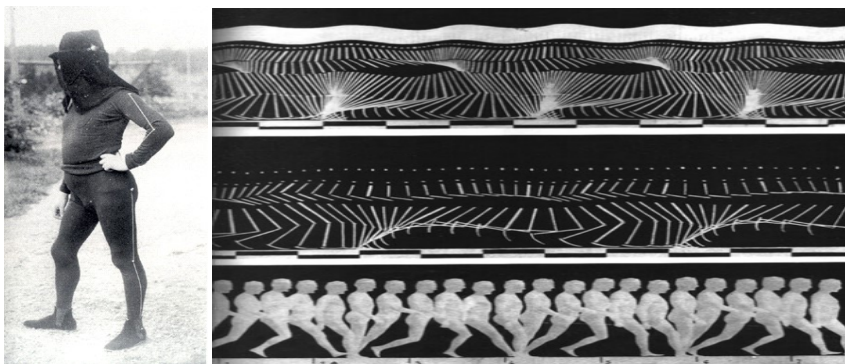


Figura 1.1 Traje especial de Marey y captura de movimiento fotografiada

Años después, cuando la animación en el cine está comenzando a surgir, Max Fleischer, un editor de arte para la revista *Popular Science Monthly*, tuvo la idea de que la animación se podía hacer con mayor facilidad y con movimientos más naturales. Se apoyó en el estilo que hoy en día llamamos animación tradicional, que para los tempranos años del siglo XX se estaba popularizando entre las audiencias y las productoras. Fleischer crea el Rotoscopio, que fue un aparato patentado en 1917, bajo el número US 1242674 A, con el título *Method of producing Moving picture cartoons* o Método para producir caricaturas de imágenes en movimiento; el método de Fleischer se basa en la idea de que los actores realizan las escenas que se desean animar, la cual luego era procesada y revelada, el film se coloca en la máquina de rotoscopio y se proyecta a la mesa de trabajo donde el animador «calca» sobre cada fotograma, para así transmitir al dibujo la naturalidad y secuencialidad de movimientos, expresiones, luces, sombras y proporciones propias de la filmación original. Este proceso sirvió de base para la animación

de escenas, cuadro por cuadro (M. Fleischer, 1917). En la figura 1.2, se pueden ver los diagramas del rotoscopio, incluidos en la petición de la patente.

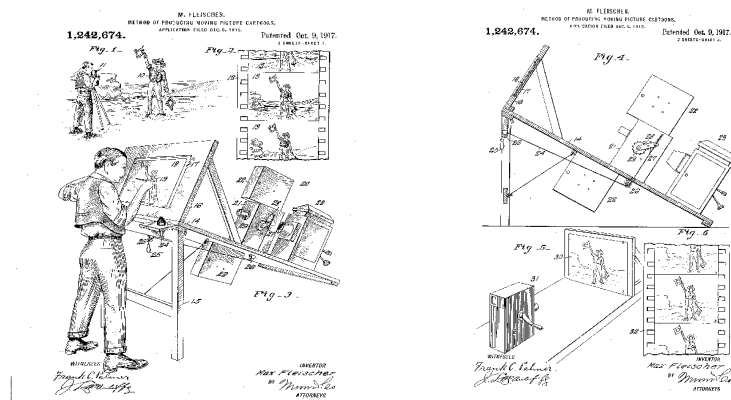


Figura 1.2 Diagramas del rotoscopio de Fleischer, incluidos en la petición de la patente (M. Fleischer, 1917)

En 1918 Max Fleischer utilizando el rotoscopio saca su primera película, llamada *Out of the Ink Well* (D. Fleischer, 1918); esta producción dio el inicio a la animación moderna. La compañía de Fleischer, antes del surgimiento de Disney, desarrolló los dibujos animados para adultos, creó personajes icónicos como Betty Boop; además de traer a las pantallas de cine a Popeye y Superman. Formó el primer estudio de animación denominado Fleischer Studio, el cual se dedicaba a la creación de cortometrajes. Luego otros estudios también incorporaron el rotoscopio en su proceso de producción de animados.

Los estudios Disney para el momento de la producción de *Blanca Nieves y los siete enanitos* (Cottrell et al., 1938), no tenían mucha gente entrenada como animadores, aunque mucho de ellos eran caricaturistas para periódicos y no estaban de acuerdo con el uso del rotoscopio, argumentaban que no tenía méritos artísticos. A pesar de ello, algunas escenas entre Blanca Nieves y el príncipe fueron realizadas utilizando el rotoscopio. De acuerdo a Kitagawa y Windsor (2008), Disney en sus producciones futuras no utilizan el rotoscopio, este se convirtió más en una herramienta para estudiar el movimiento. Debido a la hegemonía que tuvo la compañía de Disney por buena parte del siglo XX, el rotoscopio queda relegado y no se desarrolló a su máximo potencial como una herramienta para la animación. De ahí que, en la década de los 70 del siglo veinte, el rotoscopio había quedado relegado para la realización de

efectos especiales, uno de sus usos más notables fue en los sables de luz de la película de *Star Wars* (Lucas, 1977), antes de la creación de efectos generados por computadora.

Esta hegemonía general de Disney en el campo de la animación para el cine, hizo que pocas compañías en occidente utilizaran la tecnología del rotoscopio más allá de la formación del personal de animación. Los animadores de Disney se volvieron líderes en el campo, influyeron significativamente en la formación de los futuros animadores, entre los más influyentes son Frank Thomas y Ollie Johnston, quienes estuvieron involucrados en la compañía desde muy temprano. Ellos publicaron el libro *Disney Animation: The Illusion of Life*, originalmente en 1976 (Johnston & Thomas, 1995), considerados por muchos entre ellos ANW *Animation World Network* como el libro de animación más importante e influyente en el área. Es oportuno referir que estos autores forman parte del grupo que Walt Disney llamaba “*Nine Old Men*”, quienes son los maestros animadores de la época dorada de Disney. Según Menache la importancia de este libro va más allá del mundo de la animación pura, ya que los principios de la animación a que se refieren son parte importante en el proceso de la captura de movimiento, Menache recomienda entender estos principios, a fin de comprender la forma cómo se genera el movimiento animado, y cómo éste se compararía con los movimientos realistas, así como ver qué tan viables serían los movimientos que se desean capturar (Menache, 2010). Estos principios son considerados por muchos, como conocimientos básicos de la animación, y al inicio eran aplicados para la animación 2D tradicional, pero como son principios que rigen la animación misma, pueden ser también extrapolados a la animación asistida por computadora que se utiliza en la actualidad, sea esta 2D o 3D. Por esta razón, se consideró importante insertarlos a continuación.

Los principios de la animación citados por Menache son los siguientes:

- Estirar y encoger (*Squash and stretch*): El personaje pasa por cambios extremos a su figura, pero mantiene su volumen. En la animación sirve para lograr un efecto más cómico o bien más dramático. Este es el primer principio que no se puede lograr con captura de movimiento. Se puede tratar de añadir este efecto a datos capturados, pero los resultados no siempre son consistentes.

- Tiempo (*Timing*): Es el tiempo que tarda un personaje en realizar una acción. En la animación es saber determinar el tiempo de las acciones, el cual se determina con las interrupciones en el movimiento. Por esto la actuación, ya sea animada o actuada, tiene que tener el momento adecuado para transmitir la percepción necesaria, a fin de que la acción se sienta real.
- Anticipación: es una indicación de una acción por venir. Esto es típico de personajes de dibujos animados, esto guía la mirada del espectador en algunos casos, es consistente con acciones realistas, como doblar las rodillas antes de saltar. Pero no es necesariamente de la actuación humana, un buen ejecutante puede mostrar anticipación, hasta cierto punto, pero está limitado por las leyes de la física. Se verán tres pasos: anticipación (nos prepara para la acción); la acción en sí misma y la reacción (recuperación, término de la acción).
- Puesta en escena (*staging*): es la representación de una idea. Un principio de realización cinematográfica en general, el diseño de la escena y el posicionamiento de la cámara y los personajes. Poniendo en escena las posiciones claves de los personajes definiremos la naturaleza de la acción. Este principio es igualmente importantes en la animación y el rendimiento de acción en vivo.
- Acción continuada y acción superpuesta (*follow through and overlapping action*):
  - La acción continuada es la reacción que ocurre después de la acción, como la flexión de las rodillas y como se llega al suelo después de un salto de altura. En la animación 3D se utiliza mucho la acción continuada, por ejemplo, en las simulaciones dinámicas de la ropa o el pelo, lo que hacen que los movimientos capturados sean más realistas.
  - La acción superpuesta: movimientos múltiples se mezclan, se superponen e influyen en la posición del personaje. La superposición es inherente al espectáculo vivo, pero en la animación es más fácil comenzar una acción después que la otra ha terminado, lo que resulta en movimientos pausados y rígidos, poco creíbles o robóticos.
- Acción directa y pose a pose: son dos métodos de animación, que tienen sus inicios desde la animación analógica 2D, pero es también utilizada en la animación digital 2D y 3D hoy en día.

- Acción Directa (*straight ahead action*): es la animación de cuadro por cuadro de manera continua, los movimientos son más fluidos, y pueden ser más estilizados que el método de pose a pose.
- Pose a Pose: el movimiento es desglosado en una serie de poses claves, y luego los cuadros intermedios son dibujados. Originalmente, se tenían los animadores maestros que hacían las poses claves, y luego otro equipo de animadores hacen los cuadros que faltan. En la animación digital, la mayoría de los movimientos del personaje se realizan usando una variación del método de pose a pose, creando poses claves para diferentes partes, en vez de la pose para el personaje entero en un cuadro particular. Este método de animación también es conocido como *Keyframe animation*, ya que hay cuadros claves, en vez de poses completas. Esto funciona con la creación de cuadros claves y dejar que el software produzca los *in-between frames* (cuadros intermedios) por algún tipo de interpolación definida por el animador. La computadora calcula cuáles debería ser los cuadros y según los parámetros de interpolación definidos por el animador, se genera un look similar a la acción directa, aunque no estilizado. Este método es más fácil de manejar porque hay menos datos a tratar. La captura de movimiento es una especie de acción directa; como tal genera cuadros claves en cada cuadro. Esto hace que sea muy difícil de modificar. Se puede hacer que la captura de movimiento sea similar a la de pose a pose, mediante la selección de cuadros claves, eliminando los intermedios y permitiendo que la computadora haga el *Tweening*, o como se diría en español la elaboración de los cuadros intermedios.
- Entradas lentas y salidas lentas (*ease-in and ease-out*): los comienzos de los movimientos se hacen más lentos al igual que al final de los mismos, mientras se acelera a mitad del movimiento. Es un principio basado en la física del mundo real, es muy raro que un objeto entre en movimiento sin un período de aceleración, o pare de moverse sin una etapa de desaceleración. Debido a esto, se puede lograr con facilidad mediante la captura el movimiento de una actuación en directo.
- Arcos: la mayoría de las acciones no son lineales, ya que la mayoría de los seres vivos se mueven en trayectorias curvas, nunca en líneas perfectamente rectas. Esto porque al utilizar



arcos para animar los distintos movimientos del personaje, se le da una apariencia natural. Por esto cuando se está animando no se deben utilizar interpolaciones lineales entre los cuadros claves, sino que se recomienda el uso de distintos tipos de arcos.

- **Acción secundaria:** cuando se anima un personaje, se debe primero crear el movimiento principal, que suele ser el movimiento de las extremidades y la cara. Luego se crea el movimiento de otras partes del cuerpo u objetos que reaccionan al movimiento principal, como el pelo y la ropa. Estos pequeños movimientos que complementan a la acción principal, nunca deben ser más marcados que la acción principal. Los movimientos secundarios cuando se hacen con la animación tradicional (pose a pose, acción directa), representan una gran cantidad de trabajo extra. Por otra parte, la captura de movimiento es inherente a la actuación. Pero no todos los sistemas son capaces de esto, en la actualidad solo los sistemas ópticos tradicionales lo hacen, ya que se le pueden poner marcadores a la ropa para incorporarlos a la captura y sus datos puedan ser utilizados. En la actualidad los sistemas sin marcadores no son capaces de rastrear ropa.
- **Exageración:** el principio de la exageración implica acercarse o cruzar los límites de las posibilidades físicas con el fin de dramatizar la actuación del personaje. Por esto puede ser difícil hacer la captura de movimiento, ya que en momentos el movimiento puede ser imposible de realizar para tener el nivel de exageración deseado.
- **Dibujos sólidos:** es la habilidad del animador de entender al personaje, entendiendo como se mueve el personaje en un espacio tridimensional, dándoles un sentido de peso y profundidad, dándole vida al personaje. Esta es una habilidad que ha perdido un poco de importancia con la dominación de la animación por computadora, ya no se trata de poder dibujar al personaje de tal forma, sino como animar al personaje para que se perciban esas características.
- **Atractivo y personalidad (*Appeal*):** esto no tiene que ver con que sean atractivo físicamente, sino que va más en pie con el carisma de un actor. Es lo que hace al personaje interesante y que la audiencia se sienta atraída. De esta forma desarrollarle la personalidad, aunque hayan personajes que sean idénticos en su apariencia, pero pueden ser vistos de manera totalmente distinta, debido a sus diferentes personalidades, se pueden transmitir por la forma que se mueven o hablan. Se puede decir que es un principio de la actuación en general, y es por

esto que en la captura de movimiento, es esencial que se busque un actor con talento para el movimiento y que pueda dar vida a un personaje por sí solo con la forma en que se mueve, y no solamente que se busque a alguien que se pueda mover.

### 1.3.- Avances tecnológicos en la animación

El cine es el arte que más se vincula con los avances tecnológicos, pues desde sus inicios la cinematografía fue considerada en sí misma como una innovación tecnológica, ello es algo indudable, porque éste no solamente depende de las habilidades del cineasta, de su capacidad para crear ideas e imaginar escenas, sino que se requiere de un medio para plasmarlo y proyectarlo.

...el director de fotografía es quien se encarga de la parte completamente técnica, y el director responsable del guión y los actores, pero entre estos dos extremos están involucrados en la misma tarea básica: contar una historia a través de una cámara.

(Brown, 2011, p. ix)

Es así como a lo largo de la historia del cine, se ha podido apreciar una simbiosis cine y tecnología. Al respecto, Romaguera i Ramió (2007:453) indica que: “La continua evolución del cine ha dado lugar a sucesivos cambios técnicos. Algunos de ellos han sido constantes y pausados que solo con la perspectiva de los años puede advertirse la distancia recorrida”. Las Imágenes Generadas por Computadora o mejor conocido por sus siglas en inglés como *CGI* (*Computer Generated Imagery*) son un ejemplo de ello. En la actualidad se le denomina efectos visuales para distinguirlos de los efectos especiales tradicionales, que se hacían de manera mecánica con: *animatronic*, maquillaje y miniaturas. Estas imágenes juegan en la actualidad un papel trascendental como lo fue el paso del cine mudo al sonoro, del cine blanco y negro al de color. Los avances tecnológicos que ha experimentado el cine son cada vez más diversos y sofisticados, los mismos han contribuido significativamente en la producción de obras cinematográficas; estos se consideraron en un momento imposibles o poco viables de realizar, incluso hasta han tenido sus partidarios y detractores.

Los efectos visuales generados por el *CGI* son medios de expresión que han permitido exponer la idea y pensamiento del creador cinematográfico. Las imágenes generadas por Computadora se comienzan a experimentar al final de la década de los 60, con imágenes muy

básicas, debido a la poca capacidad que tenían las computadoras para procesar información en ese momento. Existen ejemplos de Filmes experimentales que utilizaron esta nueva forma de creación de imágenes, el más notable de ellos es *Hummingbird* (fig. 1.3) de Charles Csuri y James Shaffer, este film ganó un premio en la 4ta competencia Internacional de Cine experimental de Bruselas.

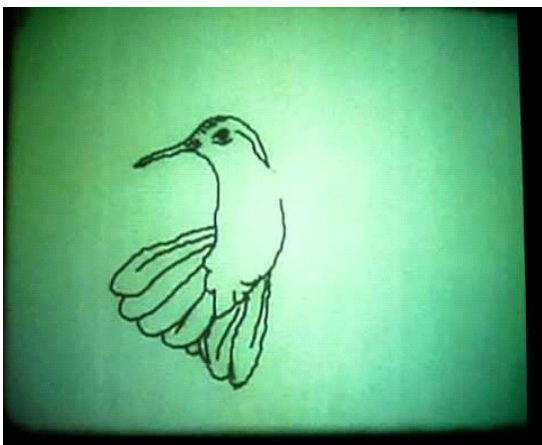


Fig. 1.3 Cuadro del Corto Hummingbird (Charles Csuri, 1968)



Fig. 1.4 Cuadro del Corto Hunger de Peter Foldes (Peter Foldes, 1974)

De la misma forma que ocurrió con la adaptación del cine a color, la adaptación al uso de imágenes generadas por computadoras, se dio mediante un proceso gradual. Los primeros pasos se dieron en las universidades de Utah y Ohio State, a principios de la década de los sesenta; allí mediante el apoyo de fondos del gobierno de Estados Unidos comienzan a desarrollar imágenes generadas en computadora, la influencia de la Universidad de Utah fue tan grande que “la mayoría de los pioneros del *CGI* estuvieron involucrados de manera directa e indirecta con su programa universitario” (Rivlin,1986); esto debido al intereses del gobierno, que otorgó fondos financieros a través de DARPA, ente encargado de promover la investigación militar. Sin embargo, estos investigadores fueron los primeros en dar el paso a la aplicación del *CGI* en las artes, así comenzaron a compartir con otras universidades que también estaban interesados en el área, de modo que ya para finales de la década de los sesenta del siglo XX, se expandió a otras universidades del mundo, entre otras están: Japón, Canadá, Europa y hasta la Unión Soviética. En este sentido, hay que reconocer que la razón principal no fue artística, sino demostrar la capacidad tanto del software como del hardware que estaban desarrollando; en consecuencia, se dieron rápidamente grandes avances y a finales de la década, ya había varios ejemplos de *CGI* para la creación artística. Mucho de estos pioneros eran gente que estaba

involucrada en el cine animado como Peter Foldes, (quien fue reconocido con un premio en Cannes y una nominación a los premios Oscar por sus cortos animados en computadora en 1974 por *Hunger* (fig. 1.4).

Para la década de los 70 del siglo XX, la animación con *CGI* comienza a dar sus primeros pasos en el Cine comercial con *Westworld*, donde se escaneó el film para aplicarle un efecto pixelado a la imagen, tratando de simular la POV del androide que persigue al protagonista, ya en la parte final se nos muestra que el androide tiene una visión infrarroja, la cual fue simulada al manipular en la computadora la imagen escaneada del film. (Ver fig. 1.5)



Fig. 1.5 *Westworld* primer uso de CGI en un largometraje comercial. ataque (Crichton, 1973)

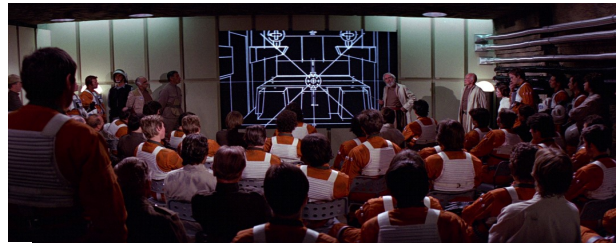


Fig.1.6 Cuadro de *Star Wars* (Lucas, 1977) Los rebeldes planean el ataque a la estrella de la muerte y los planes son una animación.

Al final de la misma década de los 70, se comienza a evidenciar los avances del poder computacional en películas como: *Star Wars* (1979) (fig.1.6), *Superman* (1978) (éste fue el primer film en tener un *Opening Credits*, hechos en computadora) y *Alien* (Scott, 1979); en esas películas se evidencia el uso de animación en 3D, con modelos de estilo *WireFrame*, debido a que las computadoras no tenían el suficiente poder de procesamiento para generar modelos sólidos en 3D. Para ese mismo momento, George Lucas funda dos compañías pioneras en el campo de la animación, utilizando *CGI*, la primera fue *Industrial Lights & Magic (ILM)* (1975), donde se trabajó tanto con efectos especiales tradicionales y luego con efectos visuales, generados por computadora. La segunda compañía que fundó fue *Lucasfilms Computer Development Division (LCDD)* (1978), dedicada exclusivamente al desarrollo de tecnología computacional para ser aplicada a la producción cinematográfica.

En el inicio de los años 80 del siglo XX comienzan a salir cortos que utilizan los primeros modelos 3D sólidos, pero no será hasta 1982 cuando se aplican a un largometraje. *Tron* fue el primer gran proyecto que sale de los estudios de Disney, el cual tuvo que ver con animación digital, contó con unos 15 minutos de *CGI*, éste fue el largometraje con el uso más extenso de

CGI en esa época y no fue sobrepasado sino hasta una década más tarde. *Tron* fue la película que enseñó los avances que había alcanzado la tecnología en un periodo relativamente corto; aproximadamente en menos de 5 años pasaron de modelos 3D estilo *WireFrame* a modelos Poligonales, con planos de colores sólidos, conocidos como *cell shading*. (ver fig.1.7). Lamentablemente, esta película no fue bien recibida por la Academia, que le negó el Oscar por mejores efectos especiales, debido a que consideraban que los efectos hechos en computadora eran como una trampa, cuando se compara con los efectos especiales tradicionales. Como bien lo señala Violet Glaze

Por desgracia ese honor no fue concedido también al director Steven Lisberger en 1982, porque la academia de las artes y las ciencias no permitió que *Tron* entrara en la categoría de Mejores Efectos Especiales, pues argumentaron que usar ordenadores para hacer una película era hacer trampas

(Glaze, 2009, p. 566)



Fig. 1.7 Cuadro de la película *Tron* durante la carrera de light cycles. (Lisberger, 1982)



Fig. 1.8 Cuadro de la película *The Abyss* donde se ve el efecto de simulación de agua en 3d. (Cameron, 1989)

A mediados de esa década, (80 del siglo XX) se ve una clara separación entre los distintos tipos de animación, *ILM* se dedica a la creación de efectos especiales foto realista que se puedan utilizar en conjunto con efectos especiales tradicionales, al mismo tiempo desarrollando la tecnología para la creación de efectos visuales a través de *Lucasfilm Computer Development Division* o *LCDD*. *LCDD* se separa de *Lucasfilm* con la inversión de Steve Jobs, quien compra la mayoría de las acciones, se convierte en *Pixar* en 1986 y se dedican a la creación de hardware para animación y modelado en 3D; así comienzan a crear cortos animados con modelos más caricaturizados para la demostración de sus productos. Luego está *Walt Disney Animation Studios* que, a pesar de haber sido pioneros en la animación 3D, cambian su enfoque a realizar películas que siguen un estilo más tradicional de animación (2D), pero en digital. Otra compañía que hizo avances en el campo de la animación con *CGI* es *Pacific Data Image*, pero ésta se dedica principalmente a créditos y trabajos para televisión. A finales de la década de los

80, siglo XX, se observa que cada vez más películas utilizan efectos visuales, generados por computadora, pero ninguna fue lo suficientemente trascendental para diferenciarse de sus predecesores; sin embargo, es por primera vez que un corto animado, completamente animado en computadora gana un premio Oscar, éste fue *Tin Toy* (1989) de Pixar. También ese mismo año se ven otros avances importantes en el campo del CGI en las películas como las siguientes: *The Abyss* (Cameron, 1989) de James Cameron, la cual ganó ese año el premio Oscar por mejores efectos especiales (ver fig.1.8) e *Indiana Jones: The last Crusade* (Spielberg, 1989) de *Steven Spielberg*, ambas películas sirvieron como base técnica y práctica en el uso de CGI para los futuros proyectos de cada director.

En el comienzo de los noventa, (sigloXX), se observa como todos esos pequeños avances de la década de los 80 terminan de establecerse como los efectos visuales que reconocemos hoy en día. La década abre con dos películas muy importantes, *Terminator 2: Judgement Day* (Cameron, 1991) de James Cameron (ver fig. 1.9), quien basándose en la tecnología simulación de agua desarrollada para la película *The Abyss* (Cameron, 1989) por *ILM*, desarrolla *Terminator 2*, como bien expresa David Jefferson al referirse a dicho film:

*Terminator 2* es probablemente la película más ambiciosa de gráficos de computadora, ya en *Tron* se apreciaron grandes efectos, pero no fue un gran éxito en la taquilla. *Tron* nunca se supuso que eran imágenes reales, pues estaban dentro de un computador. En cambio en *Terminator 2* hay personajes de acción en vivo, reales que interactúan con los gráficos por computadora. Se supone que no se debería pensar en ellos como efectos gráficos o algún tipo de cosa mística o mágica, sino que supone que pensar en ellos como algo tangible que se puede tocar

(“Visual Effects on Terminator 2 | Animator Mag - Archive,” n.d., p. 1)

Tras todas estas innovaciones, la producción cinematográfica cambio para siempre, ya que *Terminator 2* probó que el *CGI* puede lograr cosas que los métodos de efectos especiales simplemente no pueden lograr, dándole así una libertad creativa al director y al equipo de producción con herramientas, ya probadas que les permitían crear magia digital (Ver fig. 1.9). *Jurassic Park* (Spielberg, 1993), por otro lado, nos enseña que el *CGI* puede trabajar al lado de *animatronics* tradicionales y ser casi indistinguibles o diferenciar entre cuál es el real y cuál es el generado por computadora (Huls, 2013). Los efectos visuales fueron tan innovadores que



ganaron el Premio Oscar por este rubro. Ese mismo año Disney sacó *La Bella y la Bestia* (Trousdale & Wise, 1991), la cual mezcla los métodos tradicionales de animación con imágenes generadas por computadora, ello se evidencia en la escena de baile entre el personaje de Bella y la Bestia, donde el fondo fue generado por computadora para simular un efecto de dolly que es tradicional de cine no animado. (ver figura 1.10)



Fig. 1.9 Cuadro de Terminator 2, se puede ver como el efecto CGI.(Cameron, 1991)



Fig. 1. 10 Cuadro de La Bella y la Bestia en la escena del salón de baile.(Trousdale & Wise, 1991)

Para 1993 los estudios de *ILM* se encargaron de todos los efectos para la película de *Jurassic Park*, ésta es considerada por muchos como la película que volvió los efectos especiales, generados por computadora populares, no solo con Hollywood sino que también cautivó la audiencia, ya que logró lo que *Terminator 2*(Cameron, 1991) se había propuesto hacer y no lo logró; en *Jurassic Park* el público sintió tan real a los dinosaurios, como muy bien lo menciona Alexander Hulls, puesto que al referirse a estas película dijo:

Parte de esto se debe a que los espectadores aún tenían que aceptar *CGI* como algo más que una novedad. Como observó el periodista David Morgan 1993, "*El público siempre estuvo consciente de que lo que estaban viendo eran efectos especiales creados por una mano cuidadosa.*" A pesar de su éxito gracias a sus innovaciones tecnológicas *Terminator 2* no logró cautivar a la audiencia con los efectos, sino que evocan una reacción de "¿Cómo hicieron eso?". Para que los efectos sean aceptados, sus creadores deben avanzar en la tecnología hasta el punto en que la división entre la ilusión y la realidad desaparezca por completo...La revolución de *Jurassic Park* fue tecnológica (ver fig. 1.11), pero lo más importante, es que fue popular. Aunque Spielberg y Lucas hayan visto el futuro del cine en esas tomas, fue el público quien hizo ese futuro una realidad".

Part of this was because moviegoers had yet to accept CGI as anything more than a novelty. As journalist David Morgan observed in 1993, "audiences were always aware that what they were watching was carefully crafted special effects." Which is why for all of *Terminator 2: Judgment Day's* success and technological innovation, its effects didn't so much sweep audiences away as it did elicit "How did they do that?" reactions. For effects to truly break, their creators had to advance the technology to the point where the seam between illusion and reality completely disappeared.

(Huls, 2013)



Fig. 1.11 Cuadro Jurassic Park Dinosaurio realizado completamente con CGI(Spielberg, 1993)



Fig. 1.12 Cuadro de Toy Story cuando Woody conoce a Buzz sobre la cama de Andy

*Jurassic Park* ganó el premio Oscar por mejores efectos visuales y Steven Spielberg en ese mismo año ganó el premio como mejor Director y mejor Película, pero con el film *la Lista de Schindler* (Spielberg, 1994). El gran avance de la década de los 90 en el área de animación con *CGI* fue con el estreno de Pixar, la primera película que fue realizada completamente con imágenes generadas por computadora, crearon todo un nuevo estilo de películas animadas en largo metraje; la película causó un impacto tan grande, que Disney le aumenta el contrato de producción y que la joven compañía *Dreamworks* de Spielberg, compró a PDI para así crear su división de animación en 3D, que años más tarde nos trajo películas como *Shrek* (Adamson & Jenson, 2001). Esta película fue la culminación de casi dos décadas de investigación y desarrollo tecnológico en el área de CGI. La película buscó demostrar que con la animación en 3D era posible hacer cosas imposibles de hacer con técnicas tradicionales de animación. Así lo expresa Price al referirse a *Toy Story*, cuando expone que:

La primera pieza de animación, una prueba de 30 segundos, fue entregada a Disney en junio de 1992 cuando la compañía solicitó una muestra de cómo se vería la película. Lasseter quiso impresionar a Disney con una serie de cosas en la pieza de prueba que no se podía hacer en la animación tradicional, a mano, como la camisa de tela escocesa de Woody o sombras persianas venecianas que caen a través de la habitación.



(Price, 2009)

De la misma manera que *Jurassic Park*, *Toy Story* (Lasseter, 1995) no solo fue revolucionaria por lo que hizo, sino que fue una película popular, aceptada tanto por el público como por la crítica, y ese año la Academia le otorgó a Lasseter un Premio Oscar por *Special Achievement* (ver fig. 1.12); esto debido a que en ese momento no existía la categoría de mejor película animada, pero le dan el premio para reconocer el avance y la contribución que él aportó como director y a Pixar como compañía.

El éxito de *Jurassic Park*, en particular sus efectos especiales, generados por computadora marcaron el fin del Stop Motion como herramienta para la creación de efectos especiales en el cine. Phil Tippett, quien iba a ser el encargado de crear los efectos de los dinosaurios en *Jurassic Park* (ver Fig. 1.13), utilizando su técnica de Go-Motion, la cual permite que la animación Stop-Motion parezca más realista al añadir efectos de Motion-Blur. Pero después de que Spielberg vio el *Demo-Reel* de *ILM* queda convencido que los dinosaurios de CGI eran los más adecuados para su proyecto. Tras esto Tippett reaccionó diciéndole a Spielberg “I think I’m extinct” (Watercutter, 2012). De esta forma, décadas de conocimiento de animación con Stop Motion, como los maravillosos monstros de Ray Harryhausen (ver fig. 1.14) pasan a ser una herramienta que es escogida más por sus cualidades estéticas que por sus capacidades de generar efectos especiales realistas. Al punto que a pesar de no generar tanto contenido como en años anteriores, se siguen creando en la actualidad películas con esta técnica.



Fig. 1.13 Extracto de la muestra que realizó Phil Tippett a Spielberg de cómo se verían los dinosaurios para *Jurassic Park*.



Fig. 1.14 Simbad lucha con Kali en la película *The Golden Voyage of Sinbad* (Hessler, 1974) animada por Ray Harryhausen.

Desde finales de los años 90 del siglo XX, los efectos visuales generados por computadora se han vuelto casi esenciales en todas las películas, principalmente si son del género de terror, acción, ciencia-ficción. Cada vez se hacen menos frecuentes los sets enormes, porque poco a poco son remplazados por pantallas verdes. Las grandes producciones con miles de extras también han desaparecido, ya que eso se puede realizar de manera más económica por computadora. Se puede decir que, el *CGI* está cambiando en forma completa el modo en que las películas se producen, en especial con la llegada del cine digital, donde la cámara posee un sensor en vez de film, donde las moviolas se vieron remplazadas por *workstations* digitales y que el proyector ahora sea *DLP* y no 35mm. Pero esto, tiene que verse no como algo negativo, sino en forma positiva; es decir aceptando y comprendiendo que el cine se desarrolla y avanza, porque si hay algo que la historia del cine nos ha enseñado, es que el cine cambia, evoluciona con el tiempo para presentarnos las preocupaciones, deseos, aspiraciones de las generaciones que lo crean y de reflejar lo que acontece al público.

Lo antes dicho lleva a tener presente lo indicado por Mikhailovitch, quien, según Romaguera i Ramió planteaba que:

La obra cinematográfica se funda, también, no sobre una mutua limitación de ciertos medios expresivos, ni sobre la <neutralización>de unos en provecho de otros, sino en su ordenada combinación, de modo que en un momento dado puedan aparecer, en primer término, ciertos medios expresivos para significar lo más plenamente posible el elemento que conduzca del modo más directo a expresar la materia tratada, el pensamiento, el tema, la idea de la obra.

(Alsina & Romaguera, 2007, p. 497)

De ahí que, el CDI es una herramienta más en la mano del artista, que le ofrece un abanico de posibilidades para su creación cinematográfica.

#### **1.4.- Origen de la captura de movimiento:**

Tanto Menache (2010) como Kitagawa y Windsor (2008) plantean que la captura de movimiento en el campo del entretenimiento es el descendiente del rotoscopio, el cual fue adoptado por varios estudios de dibujos animados, pero pocos realmente han admitido su uso, porque muchas personas en la industria de la animación lo consideran que es hacer trampa y una profanación del arte de la animación. El rotoscopio se enfoca solo en animación en dos

dimensiones (2D), debido a que fue diseñado para los dibujos animados tradicionales, hechos a mano, con el avance de la animación tres dimensiones 3D, se provocó el nacimiento de una nueva forma de rotoscopia para 3D, que básicamente es la captura de movimiento.

En sus inicios existía diversas tecnologías para la captura de movimiento, algunas han llegado a la actualidad, estas tecnologías se han utilizado en diferentes aplicaciones, bien sea con fines médicos o militares. La captura de movimiento en gráficos por ordenador fue utilizado por primera vez a finales de la década de los 70 y principios de 1980 del siglo XX en proyectos de investigación en escuelas universitarias como la Universidad Simon Fraser, el Instituto de Tecnología de Massachusetts y el Instituto de Tecnología de Nueva York. Pero no fue utilizada en una producción real sino hasta mediados de los años 80.

In late 1984, Robert Abel appeared on a talk show and was asked if he would soon be able to counterfeit people digitally. “We are a long ways from that,” he replied. “We haven’t even figured out human motion, which is the basis, and that’s a year away.” A week and a half later, Abel received a visit on a Friday afternoon from a creative director from Ketchum, a prominent advertising agency. The visitor brought six drawings of a very sexy woman made out of chrome. She was to have Kathleen Turner’s voice and would be the spokesperson for the National Canned Food Information Council, an association formed by Heinz, Del Monte, Campbell’s, and a number of big players that sold canned food. They felt they had to make a powerful statement because the idea of buying food in cans was becoming obsolete, so they wanted to do something really different and outrageous, and they wanted it to air during the Super Bowl in January 1985. “Can you do it?” asked the client. “You’re certainly here a lot earlier than I would have planned,” replied Abel, and asked the client to wait until the end of the weekend for an answer.

(Menache, 2010, p. 4)

Menache (2010) relata específicamente lo expresado por Robert Abel de cómo fue abordado por un representante del *National Canned Food Information Council*, a fin de hacer un comercial destinado a salir al aire en el *Superball* de Enero de 1985; su idea de publicidad era tener específicamente una mujer atractiva, hecha de cromo que dijera que: todavía para el año 3.000 la comida enlatada era la mejor forma de preservar alimentos a largo plazo. Abel anonadado dijo: no me esperaba que me vinieran a pedir este tipo de cosas tan pronto y le pide al mencionado representante que le diera el fin de semana para ver si eso era posible hacerlo y le daría la respuesta. Es importante tener en cuenta que para esta época las capacidades gráficas

de las computadoras disponibles en el mercado, no iban más allá de la creación de logotipos en movimiento, paisajes, y objetos sólidos. Es decir, no existían la cantidad de software para la animación que existe hoy en día, justo en ese momento las compañías que se dedicaban a la animación creaban su propio software para su uso interno, no podían saber si ese programa era capaz de hacer lo que se pedía, sin hacer las pruebas previas. Por esta razón, Abel y su equipo se encerraron ese fin de semana completo para ver si podrían realizarlo en el plazo de tiempo que el cliente requería, el cual era apenas ocho semanas.

En ese momento, no se podían hacer métodos de animación moderno como *Keyframe*, ya que las computadoras de la época no tenían la capacidad para procesar ese tipo de movimiento, entonces decidieron ir con la captura de movimiento, pero en lugar de tener un programa con el algoritmo que era capaz de rastrear los puntos por sí mismo, ellos tenían que rastrear el punto cuadro por cuadro, por esta razón ellos crearon este algoritmo tras una prueba interna que hicieron durante ese fin de semana, donde un miembro del equipo con solamente interiores puestos, se le puso tape negro eléctrico en las articulaciones, para que tuviera la función de marcadores. Mientras él realizaba movimientos lentos, se le iba fotografiado con un cámara polaroid, luego las fotos se escanearon y procesaron para lograr así crear las base del algoritmo que le permitiría realizar la captura de movimiento, la cual tenía sus limitaciones, en relación a la captura de movimiento moderno que tiene la capacidad de rastrear los marcadores a lo largo de la toma, sin tener que ser fijados por el operador. En este caso, los marcadores tuvieron que ser fijados a mano por el operador de la computadora, cuadro por cuadro. Y tras calcular cuánto sería el tiempo para digitalizar la data capturada, se dieron cuenta que podían hacer el comercial en el tiempo requerido por el cliente.

La agencia de la publicidad consiguió una joven bailarina como modelo, a ella se le puso una serie de marcadores a lo largo de los dieciocho puntos de las articulaciones, pero en lugar de utilizar tape negro, usaron marcador negro y una silla que pudiera rotar 360° para que así la modelo pudiera realizar los movimientos, sin tener ningún obstáculo que impidiera el movimiento o que ocultara algún punto del marcador, mientras se fotografiaba en diferentes ángulos. El procesamiento de la data capturada se demoró cuatro semanas y media, usando varias computadoras.

Mientras se realizaba el procesamiento de la data, crearon el *wireframe* del modelo de manera separada para que cada parte fuera puesta junta y se les aplicó la data procesada del movimiento, allí surgió un nuevo problema, que era cubrir el *wireframe* con una textura cromada. Si se utilizaba el método común para la época que era *ray-tracing*, sería imposible realizarlo en el tiempo deseado sin que se utilizara una computadora multimillonaria. Charlie Gibson ideó una forma de resolver el problema en la cual la textura era animada, siguiendo la topología del cuerpo del personaje a medida que éste se moviera, creando lo que hoy en día se conoce como *reflection map*. Lo último que tenían que lograr en las dos semanas que les quedaba era renderizar la animación, esto lo pudieron lograr porque pidieron ayuda a individuos dispersos a lo largo de Estados Unidos, quienes le dieron acceso a un total de 60 computadoras del mismo modelo de las que ellos tenían, para poder realizar la renderización, algunas de estas computadoras estaban en lugares tan lejanos como Hawái, Alaska, Florida y algunos lugares de Canadá. Gracias a esta ayuda fueron capaces de entregar el spot publicitario dos días antes de la fecha pautada. Este Spot marcó el momento del primer personaje con captura de movimiento que se utilizó con fines comerciales, a pesar de haberse llamado este personaje *Brilliance*, el mismo fue conocido en la industria de la animación como *sexy robot*.

Para Menache (2010) en el año 1988 se dieron grandes avances tecnológicos en la captura de movimiento de manera pública, entre ellos están los avances de Graf y Wahrman y los de *Pacific Data Image (PDI)*, quienes se alejaron de la captura de movimiento óptica anteriormente descrita, para pasar a la captura de movimiento electromagnética que permitía su realización en tiempo real. Específicamente en 1988, comenzó la colaboración de *PDI*, con *Jim Henson and Associates*, a fin de lograr la creación del personaje de *Waldo C. Graphic* de la serie de televisión *The Jim Henson Hour*. (Ver figura 1.13) Una de las particularidades de este proyecto, además de ser uno de los primeros personajes hecho con captura de movimiento que forma parte de un programa de televisión, fue la necesidad de que hubiera una representación digital renderizada en tiempo real; esto se debe a que las actuaciones del show dependían de la espontaneidad y las reacciones realizadas, tanto por los actores como de los títeres en escena. Aunque la renderización en tiempo real no es la definitiva, la misma se realiza después de la grabación para mejorar la calidad del producto final, de modo que así pudiera salir al aire.

Como se dijo anteriormente, esto solo era posible si se realizaba con sensores como los potenciómetros, los cuales se basaban en el sistema inalámbrico para el control telemétrico de un robot; es importante señalar que este sistema fue el utilizado por la compañía de los Mopets de Henson para la animación de la rana Kermit en una escena, donde está montando en una bicicleta y el público vio el cuerpo entero del personaje en la toma, mientras el mueve la cabeza y canta una canción. El robot de Kermit era controlado por Jim a través de un artefacto de goma espuma que imitaba la cabeza y parte del cuello de Kermit, el cual funcionaba con algunos potenciómetros conectados a una señal de radio, que transmitían los movimientos realizados por Jim y que eran imitados por el robot.



Fig. 1.15 a. El personaje Waldo C. Graphic con su renderización foto realista, no la versión en vivo,

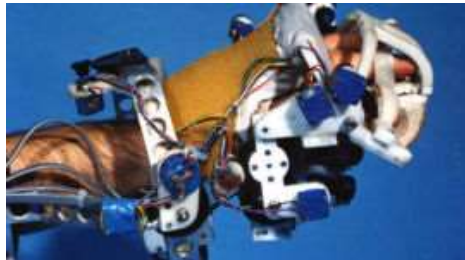


Fig. 1.15 a. Cabecial del sistema de captura de movimiento de waldo sin la capa protectora, de goma espuma



Fig. 1.15 c Jim Henson, en su programa de television explicando como funciona el sistema de captura de movimiento de waldo.

En el caso de Waldo, *PDI* (ver figura 1,15) construyó un brazo mecánico para poder capturar la posición del personaje en el espacio, y en vez de tener un transmisor, los sensores se conectaban directamente a una computadora que grababa el movimiento. Esto permitía que el titiritero se pudiera mover por el espacio y hacer movimientos que no eran posibles con títeres tradicionales, ya que su cuerpo sería removido de la imagen digitalmente. En algunos casos, se complementaban los movimientos con pequeñas animaciones secundarias, las cuales eran imposibles de capturar, como movimientos del vientre, y transformaciones en otros objetos. El programa tenía un formato semanal y el personaje de Waldo tenía una aparición de aproximadamente un minuto, lo cual dejaba 5 días para la producción, la cual era constante. Rex Gignron, quien era el responsable de la captura del personaje, dice que hasta la actualidad piensa que la creación de ese personaje con captura de movimiento es una de las mejores con las que había estado involucrado, además que era impresionante como los titiriteros le dan vida a los títeres.

Basándose en el trabajo que se hizo en el personaje de Waldo, todavía en 1988, *PDI* busca crear un dispositivo de la parte superior del cuerpo (cabeza, torso y brazos). Llamaron a este dispositivo mecánico "exoesqueleto"; se basa en potenciómetros ópticos en cada articulación, hubo dos iteraciones del exoesqueleto, la primera era analógica, pero era ruidosa e imprecisa. El segundo exoesqueleto, que se utilizó para las películas de Barry Levinson "Toys", en una secuencia donde se necesitaba que pareciera que alguien tenía visión de rayos X y veía una reunión en curso, Jamie Dixon que era el supervisor de Efectos de *PDI* y el encargado del proyecto terminó siendo el intérprete de los personajes en escena. Esta captura tuvo algunos problemas técnicos en la recolección de datos, los cuales necesitaron ser limpiados después de la actuación, pero se disponía de datos lo suficientemente buenos para poder utilizarse en una cantidad de tiempo razonable. Por lo cual, *Toys* fue la primera película en la que un personaje digital se ha creado con éxito, utilizando la captura de movimiento.

Más adelante *PDI*, reemplaza los potenciómetros en sus trajes, con sensores electromagnéticos, del tipo *Flock of Birds*, manufacturados por *Ascension Technology Corporation*, este exoesqueleto fue implementado para varios proyectos, uno de ellos era un *Pilot* para un show llamado *The Late Jackie Lenny* para *Comedy Central* en la que un esqueleto tenía un programa de entrevistas, allí se entrevistaría a cómics en vivo o haría un acto de *stand-up*. Para todos los proyectos relacionados con el exoesqueleto electromagnético, *PDI* fabrica una serie de accesorios, destinados a mantener los sensores en su lugar y hacer más fácil la experiencia de captura para los actores. La mayoría se basan en tiras de velcro para sostener los sensores. Mientras que el sensor que representaba la cabeza fue colocado en la parte superior de una gorra de béisbol, de modo que cuando el intérprete se quitaba la gorra, el personaje se vería como si se hubiera quitado la cabeza. Este sistema de captura de movimiento a base de sensores electromagnéticos tenía algunos problemas con la interferencia, el cual era muy evidente, porque el estudio de captura se encontraba en el tercer piso de un edificio de tres pisos y los equipos de aire acondicionado se encontraban en el techo, justo encima del escenario. Así que si el aire acondicionado se encendía mientras se estaba utilizando el sistema, los personajes parecían como si estuvieran temblando de arriba hacia abajo en la pantalla.

Por su parte, la compañía de deGraf y Wahrman, según Menache (2010) fue contratada por la compañía *Silicon Graphics* para que hicieran la demostración de sus nuevos modelos de

computadoras, conocidas como las 4D en la *SIGGRAPH'88*. Específicamente, *Silicon Graphics* quería que en esa convención se diera una interactividad entre un personaje animado y la audiencia. Esto lo lograron haciendo una especie de títere electromagnético, el cual era controlado por el operador, estos movimientos eran traducidos y renderizados por la computadora en tiempo real, creando así la ilusión que el personaje animado estuviera interactuando realmente con la audiencia. Esta experiencia dio base para que Brad deGraf, tras la disolución de la que tenía con Wahrman, fuera a trabajar con *Colossal Picture*, en donde reescribe el software de captura de movimiento, basado en marcadores electromagnéticos en tiempo real, con el propósito de poder especializarse en la presentación de soluciones en captura de movimiento en tiempo real para programas de televisión. Un ejemplo de esto fue el personaje de *Moxy* para el canal de *Cartoon Network* en 1993.

Ya para los años 90 más y más compañías estaban ofreciendo servicios de captura de movimiento, éstas funcionaban como intermediario entre los que manufacturaban los equipos para la captura de movimiento, su software original y los que necesitaban de este tipo de servicio. Es importante señalar que, para ese momento las únicas empresas que manufacturaban este tipo de equipos lo hacían para el área de la medicina, algunos departamentos de ciencias de universidades y para el ejército, pero no para compañías del entretenimiento. La mayoría de sistemas de captura de movimientos que se ofrecían eran los ópticos, porque los sistemas electromagnéticos se realizan adaptados a los intereses del cliente.

Entre las compañías más importantes que ofrecían su servicio al área de entretenimiento en la década de los 90, estaban Biovision en San Francisco y *TSi* de Los Ángeles, estos sistemas eran altamente costosos y difíciles de operar, necesitaban un alto conocimiento técnico, debido a que era muy complejo la captura de data que se pudiera usar y también era casi imposible verificar si la data era posible de utilizarla durante la captura, porque requerían computadoras de alto rendimiento; además que tomaba horas hacerlo y estas compañías se especializaban en dar este servicio, facilitando de esa forma el acceso a la captura de movimiento. Los principales clientes de estas compañías fueron los de la industria del videojuego, ya que las consolas se volvieron más poderosas y necesitaban que los movimientos fueran más realistas. Para principios del siglo XXI con los avances tecnológicos, los sistemas de captura de movimiento se vuelven más accesibles, tanto en costo como en capacidad técnica, ya que las



manufacturadoras de los sistemas de captura de movimiento buscaron ampliar su mercado, al introducir un Plug-in que era compatible con los software de animación disponibles en el mercado, a objeto de habilitarlos para hacer la captura de movimiento. Esto llevo al surgimiento de más compañías que ofrecían servicios de captura de movimiento alrededor del mundo, y a su vez hizo que las grandes productoras de videojuegos, como *Electronic Arts* crearan su propia división de captura de movimiento para ofrecer este servicio desde la misma compañía sin intermediarios. En los actuales momentos existen diversos sistemas de captura de movimiento, estos han ido desarrollándose cada vez más.

### **1.5.- El proceso de captura de movimiento:**

Este es el tema central del presente trabajo, el uso de la captura de movimiento en la animación, ésta se diferencia de la animación digital, porque la captura de movimiento es basada en la realidad y naturaleza humana del actor, en cambio la animación digital surge de la imaginación del artista que la produce, por ello es imprescindible definir primero ¿Qué es la captura de movimiento? Menache (2010:2) la define de la siguiente manera:

Motion capture is the process of recording a live motion event and translating it into usable mathematical terms by tracking a number of key points in space over time and combining them to obtain a single three-dimensional (3D) representation of the performance. In brief, it is the technology that enables the process of translating a live performance into a digital performance. The captured subject could be anything that exists in the real world and has motion; the key points are the areas that best represent the motion of the subject's different moving parts. These points should be pivot points or connections between rigid parts of the subject. For a human, for example, some of the key points are the joints that act as pivot points and connections for the bones. The location of each of these points is identified by one or more sensors, markers, or potentiometers that are placed on the subject and that serve, in one way or another, as conduits of information to the main collection device. From now on, when speaking generally about these, I will refer to them as "markers."

(Menache, 2010, p. 2)

Basándonos en lo establecido por Menache, la captura de movimiento se definirá en este trabajo como: El proceso de grabar el movimiento en vivo de un actor, animal u objeto, con el propósito de traducir dicho movimiento a datos matemáticos que puedan ser utilizados para rastrear puntos claves (marcadores) o siluetas (sin marcadores), en un espacio tridimensional y en un periodo de tiempo determinado. La captura de movimiento siempre tendrá por fin el crear

una representación digital de los movimientos de la actuación real, para que estos sean aplicados a un modelo 3D que sirva como representación virtual o digital del actor. Este proceso se puede equiparar con la técnica de rotos-copio utilizado en la animación 2D, porque facilita la creación de animaciones con movimientos realistas de cualquier objeto que se pueda conseguir en la realidad. De igual forma, la captura de movimiento se puede utilizar para grabar el movimiento de cualquier sujeto que tenga movimiento propio y aplicarlo a su homólogo digital.

Es importante aclarar que, en este trabajo se enfocó más en el uso de la captura de movimiento para la grabación de sujetos con movimientos humanoides, ya que la mayoría de los avances en los últimos años se han encaminado en esta área de la captura de movimiento. Existen diversos mecanismos y sistemas para la captura de movimiento, sin embargo todos ellos se basan en los mismos principios. Se trata de rastrear los puntos claves del sujeto, cuyo movimiento se desea grabar, estos representan las diferentes áreas de movimiento del sujeto, son los puntos de pivotes que conectan las partes rígidas del sujeto. Por ejemplo, si el sujeto es un ser humano, los puntos pivotes son las articulaciones (codos, rodillas, hombros, etc.) las cuales sirven de conectores entre los huesos. Cada una de las articulaciones que sirven como puntos de referencia son marcadas con sensores, reflectores u otros conductores de información, los cuales remiten la información a la computadora u otro aparato para que sea recolectada. Estos puntos de referencia o pivotes son conocidos como marcadores o *markers*, debido a que la mayoría de la tecnología fue desarrollada en países angloparlantes.

### **1.5. 1.- Tipos de Sistemas de captura de movimiento**

Existe una gran variedad de tecnologías disponibles para la captura de movimiento, desde los sistemas ópticos y mecánicos, que fueron los primeros sistemas desarrollados, hasta los sistemas magnéticos y de radio que han buscado ampliar el uso de la captura de movimiento más allá de la medicina, ingeniería y animación. Cada tipo de sistema tiene sus ventajas y desventajas, dependiendo del área dentro del cual se aplique la captura de movimiento. Estos sistemas se pueden clasificar en:

- Sistemas ópticos
- Sistemas no ópticos.

### 1.5.1.1 - Sistemas Ópticos

Los sistemas ópticos son los más antiguos de todos los sistemas de captura de movimiento, según Kitagawa y Wdsor (2008), se puede decir que es el sistema precursor, ya que para finales del siglo XIX era la única tecnología disponible que podía capturar el movimiento de un ser vivo. Todos los sistemas ópticos modernos continúan basándose en el mismo principio; es decir, la simple utilización de data capturada a través de sensores de imagen, para triangular la posición del sujeto en un espacio 3D, relativo a la posición de dos o más cámaras que hayan sido previamente calibradas para proporcionar proyecciones superpuestas. Estos primeros sistemas ópticos surgieron en los laboratorios de las universidades estadounidenses, financiados principalmente por el departamento de defensa de ese país. La mayoría de estos sistemas fueron diseñados y fabricados para aplicaciones médicas y militares; por esta razón, carecían de las características importantes para las aplicaciones gráficas de computadora. Fue apenas a finales de la década de los 70 del siglo XX cuando se logró el paso de los sistemas analógicos, desarrollados a finales del siglo XIX a los sistemas digitales, que se conocen en la actualidad.

Según Menache el primer sistema óptico digital diseñado con gráficos por computadora fue el *Vicon 8*. Este fue el primero capaz de soportar código de tiempo SMPTE, ésta es una marca de tiempo utilizado por la mayoría de las aplicaciones en el cine y la televisión. Antes de este sistema no había una forma fácil de combinar el video con los datos de movimiento reales. Tener el código de tiempo en los datos de movimiento permitió editar los archivos de movimiento, como se haría con el vídeo en vivo (*Live-Action*), y planificar adecuadamente la asociación de los personajes con sus fondos respectivos. Otra característica muy útil del sistema *Vicon 8* es que su software introdujo la grabación de videos en el formato AVI, al mismo tiempo de la captura de movimiento, estos archivos de película sirven de referencia para el post-procesamiento y aplicación de datos capturados.

De acuerdo a Menache, en la actualidad los sistemas ópticos son métodos muy precisos para la captura de ciertos movimientos, especialmente cuando se utiliza un sistema de tecnología de última generación. A pesar de todos los avances que se han generado en los últimos 30 años, en la mayoría de los casos, la captura de movimiento todavía no es un proceso que se realiza en tiempo real; una respuesta inmediata no es posible al menos que los

movimientos no sean demasiado complejos y no haya demasiados personajes en captura. Los datos adquiridos ópticamente pueden requerir un amplio uso de *post-processing* para ser utilizables, debido a esto los costos de operación pueden ser muy altos (Menache, 2010).

Según Menache, la mayoría de los sistemas ópticos disponibles se basan en un único equipo que controla la entrada de varias cámaras, ésta se denomina con las siglas *CCD* en inglés (*charge- coupled device*). Los *CCDs* son un sensor de cámaras digitales sensibles a la luz que utilizan una matriz de células fotoeléctricas (también llamados píxeles) para captar la luz, y luego medir la intensidad de la luz para cada una de las células y la creación de una representación digital de la imagen. Una cámara *CCD* contiene una matriz de píxeles que pueden variar en la resolución de tan bajo como 128 x 128, en la actualidad pasan ya de los millones de píxeles. El estado actual es de 41 millones de píxeles (41 megapíxeles) en cámaras fotográficas y unos 16 megapíxeles en cámaras de video, como las que se utilizan para la captura de movimiento, y es importante aclarar que ese número va en continuo aumento cada año. Obviamente, cuanto mayor sea la resolución, mejor será la captura, pero hay otras características que se deben tomar en cuenta. La tasa de muestras por segundo, o frecuencia de imagen tiene que ser lo suficientemente rápida para capturar los matices de movimientos muy rápidos. En la mayoría de los casos, 60 muestras por segundo son más que suficientes, pero si los movimientos son muy rápidos, como un campo de béisbol, se requieren muchas más muestras por segundo. Para los estándares de hoy en día, una cámara *CCD* con una resolución de 16 megapíxeles sería capaz de producir hasta 240 muestras por segundo en esa resolución. Sin embargo, para capturar movimientos muy rápidos, la resolución tendría que ser reducida para evitar que el sensor se sobre caliente y que el procesador de imágenes pueda seguir el paso de las imágenes (Menache, 2010).

#### 1.5.1.1.1.- Funcionamiento de los sistemas ópticos:

Kitagawa y Windsor plantean que todo sistema óptico debe calibrarse, a fin de que funcionen adecuadamente y pueda rastrear a los actores. Esto se logra cuando se tienen todas las cámaras posicionadas para rastrear un objeto con dimensiones conocidas y que el software puede reconocer, tales como un cubo o una varita con marcadores. Mediante la combinación de las vistas de todas las cámaras del calibrador (objeto de dimensiones conocidas), el software puede determinar la posición exacta de cada cámara en el espacio. Si por lo menos una cámara

es movida o golpeada ligeramente, se debe realizar una nueva calibración, ya que el rastreo no cuadraría con la calibración previa. Por eso, se recomienda calibrar el sistema varias veces durante una sesión de captura, especialmente si el estudio está situado en un terreno inestable o cerca de una vía muy transitada, ya que cualquier tipo de movimiento o vibración puede cambiar la posición de una cámara (Kitagawa & Windsor, 2008).

Igualmente, Kitagawa y Windsor plantean que, se necesitan al menos dos puntos de vista ya calibrados para rastrear la posición de un punto 3D, pero se necesitan también cámaras adicionales para evitar la oclusión y mantener una línea de visión directa con al menos dos cámaras para cada marcador. Eso no quiere decir que, más cámaras son mejores, ya que cada cámara adicional aumenta el tiempo de post-procesamiento, siendo a veces poco rentable; por eso es recomendable el uso estratégico de cámaras para hacer que las sesiones de captura sean lo más eficientes posibles (Kitagawa & Windsor, 2008).

Menache plantea que, una vez que las vistas de cámara están digitalizadas en el ordenador, se comienza el momento del post-procesamiento, procurando seguir estos pasos:

1. El primer paso es que el software trate de producir una reproducción limpia de sólo los marcadores. Hay diferentes métodos de procesamiento de imágenes que se utilizan para minimizar el ruido y aislar los marcadores, separándolos del resto del medio ambiente. El enfoque básico es separar todos los grupos de píxeles que exceden un umbral de luminosidad predeterminado. Si el software es lo suficientemente inteligente, utilizará cuadros previos o consecutivos para ayudar a resolver cualquier cuadro particular.
2. El segundo paso es determinar las coordenadas 2D de cada marcador para la vista de cada cámara. Estos datos son utilizados en combinación con las coordenadas del resto de cámaras para obtener las coordenadas en 3D de cada marcador.
3. El tercer paso es identificar en cada marcador a lo largo de una secuencia. Esta etapa requiere de la mayor asistencia por parte del operador, ya que la asignación inicial de cada marcador tiene que ser registrada manualmente. Después de esta asignación, el software intenta resolver el resto de la secuencia hasta que pierda de vista a un marcador debido a la oclusión, momento en el cual el operador debe reasignar los marcadores en cuestión y continuar el proceso. Este proceso continúa hasta que se resuelva toda la

secuencia y se guarda un archivo que contiene los datos de posición de todos los marcadores.

El archivo producido por este proceso, según Menache contiene una secuencia de posiciones globales de los marcadores a lo largo del tiempo, lo que significa que sólo las coordenadas cartesianas (x, y, z) de cada marcador por cada cuadro, no incluye la jerarquía de las extremidades o rotaciones (Menache, 2010). Es posible utilizar este archivo para la animación en la computadora, pero se requiere una configuración más extensa dentro del software de animación con el fin de resolver el esqueleto final para ser utilizado. Los directores técnicos experimentados pueden beneficiarse mediante el uso de los datos de este archivo, ya que permite un mayor control sobre lo que se puede hacer en la preparación del personaje. Para el usuario promedio, sin embargo, los datos deben ser procesados un poco más, al menos hasta el punto de incluir una jerarquía esquelética con las rotaciones de las extremidades. La mayoría de los sistemas ópticos incluyen sistemas de datos de edición que permiten al usuario producir los datos jerárquicos de rotación antes de importarlo al software de animación.

La adquisición de datos se lleva a cabo tradicionalmente, usando marcadores especiales, colocados a un actor; sin embargo, de acuerdo a lo planteado por Menache los sistemas más recientes son capaces de generar datos precisos mediante el seguimiento de características de la superficie, identificados dinámicamente para cada sujeto particular; es decir, se pueden realizar con un sistema sin marcadores. Estos sistemas producen datos con 3 grados de libertad por cada marcador, y la información de rotación debe inferirse a la orientación relativa de tres o más marcadores; por ejemplo, el hombro, el codo y la muñeca son marcadores proporcionando el ángulo del codo. El seguimiento de un gran número de actores o la ampliación de la zona de captura, se logra mediante la adición de más cámaras.

Existen diferentes sistemas ópticos, estos pueden clasificarse en:

1. Sistemas con marcadores:
  - a. Pasivos:
  - b. Activos
2. Sistemas sin marcadores:
  - a. Sistemas con múltiples cámaras.
  - b. Sistemas con sensores de profundidad.

#### 1.5.1.1.2.- Los sistemas ópticos con marcadores pasivos

Los sistemas ópticos pasivos son los más comunes, ya que es el más antiguo de los tipos de captura de movimiento, el número de cámaras empleadas por lo general no son menos de 8 y no más de 32, pero hay casos en los que se utilizan cientos de cámaras, como se explicará más adelante. Según Menache se captura la posición de los marcadores refractivos a velocidades entre 30 y 2000 muestras por segundo. Las cámaras están normalmente equipadas con sus propias fuentes de luz que crean una reflexión direccional de los marcadores, las cuales generalmente son esferas cubiertas con un material, como por ejemplo, una cinta refractiva. Las fuentes de luz que se prefiere utilizar son la roja, porque crea menos distorsión visual para el usuario. Los infrarrojos también se usan, pero es un poco menos eficaz que el rojo visible. Las esferas utilizadas como marcador puede variar de diámetro, desde unos pocos milímetros de diámetro para las capturas de área facial y pequeñas, a un par de centímetros para otras áreas. Otra característica importante es la sincronización del obturador, la velocidad del obturador de la cámara se puede sincronizar con fuentes externas, como los diodos emisores de luz (*LEDs* por sus siglas en inglés) con los que las cámaras de captura de movimiento óptico suelen ser equipados.

Marcadores pasivos del sistema óptico están recubiertos con un material retro-reflectante para reflejar la luz que se genera cerca de la lente de las cámaras. Se puede ajustar al lente de la cámara, de modo que sólo los marcadores reflectantes brillantes se muestrea, haciendo caso omiso de la piel y tela (Ver Fig. 1.16). Un objeto con marcadores en posiciones conocidas se utiliza para calibrar las cámaras, obtener sus posiciones, y medir la distorsión del lente de cada cámara, para así obtener la mayor precisión posible. Si dos cámaras calibradas ven un marcador, se puede obtener una solución tridimensional. Normalmente, un sistema con marcadores pasivos constará alrededor de 2 a 48 cámaras. Existen sistemas que tienen más de 300 cámaras para tratar de reducir el intercambio de marcador. Se requieren cámaras adicionales para la cobertura completa alrededor de la captura de sujetos y múltiples sujetos.



Fig. 1.16 Un traje de captura de movimiento común, con sus marcadores pasivos refractantes tradicionales.



Fig. 1.17 El actor Andy serkis en el set de Dawn of the Planet of the Apes (Reeves, 2014) con un traje de captura de movimiento diseñado para ser utilizado en locacion en vez de un estudio, funciona con marcadores pasivos.

A diferencia de los sistemas de marcadores activos y sistemas magnéticos, de acuerdo a lo planteado Menache los sistemas pasivos no requieren que el usuario use cables o equipos electrónicos. En lugar de ello, cientos de pelotitas de goma se unen con cinta reflectante, que debe ser reemplazado periódicamente. Los marcadores se suelen conectar directamente a la piel (como en la biomecánica), o están con velcro en un actor que lleva un traje de spandex / lycra de cuerpo completo, diseñado especialmente para la captura de movimiento (Ver Fig. 1.17.). Este tipo de sistema puede capturar un gran número de marcadores a velocidades de cuadro generalmente alrededor de 120 a 160 fps, aunque mediante la reducción de la resolución y seguimiento de una región más pequeña de interés pueden realizar un seguimiento tan alto como 10.000 fps.

#### 1.5.1.1.3.- Los sistemas ópticos con marcadores activos

Los sistemas activos, utilizan marcadores especiales que requieren *LEDs* con su propia fuente de poder para generar luz (ver fig. 1.18). Son triangulados de manera similar a los sistemas pasivos en el cual la computadora interpreta la posición relativa entre los marcadores para calcular el movimiento. La ventaja de que los marcadores generen su propia luz, en vez de reflejar luz emitida por una fuente externa es la ley cuadrática inversa, en la cual se establece que una fuente puntual en todas direcciones por igual, la intensidad de la misma disminuye, de acuerdo con el cuadrado de la distancia a la fuente de emisión, en la captura de movimiento significa, ya que el sistema no refleja luz sino la emite, un sistema activo puede tener una mayor área de posible captura; además de mejorar la calidad de la data capturada, ya que la luz tiene menos interferencia.





Fig. 1.18 Sistema de captura de movimiento activo profesional creado por la compañía Pshase Space

Esta tecnología fue utilizada en películas como *Van Helsing* (Sommers, 2004), donde utilizaron un sistema óptico activo para la escena del vuelo de las esposas de Drácula, en un set que era particularmente grande en comparación con sets de captura de movimiento convencionales. En la película *Rise of the Planets of the Apes* (Wyatt, 2011), el estudio Weta utilizó un sistema óptico activo por razones similares, pero además de esto cada marcador tenía una sincronización especial con el software de captura de movimiento, lo cual permitía reconocer cada marcador de manera individual sin tener que recurrir a ecuaciones algorítmicas que requerirían procesamiento extra. Esto acelera el procesamiento de la data cuando la misma se utilice en aplicaciones que requieran captura en tiempo real.

Un aspecto importante a considerar en la selección de estos sistemas con marcadores, es conocer cuáles son sus ventajas y desventajas, en la tabla 1.1 se presentan las ventajas y desventajas de los sistemas ópticos con marcadores.

Tabla 1.1: Ventajas y Desventajas de los sistemas ópticos con marcadores

Ventajas	Desventajas
----------	-------------

<ul style="list-style-type: none"> <li>• Los datos ópticos pueden ser extremadamente precisos.</li> <li>• Un mayor número de marcadores puede ser rastreado.</li> <li>• Es fácil de cambiar configuraciones de los marcadores.</li> <li>• Es posible obtener aproximaciones a esqueletos internos mediante el uso de grupos de marcadores.</li> <li>• Ejecutantes no están limitados por los cables.</li> <li>• Los sistemas ópticos permiten un área de actuación más grande que la mayoría de los otros sistemas de hoy en día.</li> <li>• Los sistemas ópticos tienen una frecuencia de muestreo más alto de captura, lo que resulta en más mediciones por segundo.</li> </ul>	<ul style="list-style-type: none"> <li>• Datos ópticos requiere una amplia post-processing, por lo que los costos de operación son altos.</li> <li>• El hardware es caro. Un sistema óptico de nivel de entrada puede costar más de \$ 50.000 y un sistema de gama alta puede costar más de un millón.</li> <li>• Los sistemas ópticos no pueden captar movimientos cuando los marcadores están ocluidos durante un largo periodo de tiempo.</li> <li>• Capturar debe llevarse a cabo en un ambiente controlado, lejos de la luz amarilla y el ruido reflectante.</li> </ul>
---	--

Fuente: Menache (2011) Kitagawa y Windsor (2008)

#### 1.5.1.1.4.- Los sistemas ópticos sin marcadores:

Los sistemas sin marcadores funcionan con los mismos principios que el resto de los sistemas ópticos que utilizan marcadores, ya que todos los sistemas ópticos se basan en los mismos principios matemáticos de la triangulación, que utiliza la trigonometría de los triángulos para determinar posiciones de puntos o medidas de distancia. La diferencia principal se encuentra en que los sistemas con marcadores, se enfocan en un solo punto, el cual es el marcador. Mientras que los sistemas sin marcadores van más allá, de rastrear los marcadores. Van hacia la línea del desarrollo de la visión computarizada, que no es solamente el poder rastrear una bola de luz en un volumen de captura, porque un sistema óptico sin marcadores necesita para poder funcionar, por lo menos, de alguna forma rudimentaria de saber qué es lo que ve; un ejemplo claro de esto son los carros autónomos que superan el sistema óptico tradicional, sea este pasivo o activo. Se puede decir que el sistema óptico sin marcadores es uno de los más recientes para cumplir el objetivo de capturar y procesar la data generada por el movimiento, a fin de poder aplicarlo en la animación. Los sistemas sin marcadores han sido teorizados desde finales del siglo XX, como Yonemoto y su equipo utilizaron múltiples computadoras interconectadas para realizar captura de movimiento a través de múltiples cámaras, de la misma forma que algunas universidades alrededor del mundo han hecho experimentos en el área durante la primera década del siglo (Yonemoto, Matsumoto, Arita, & Taniguchi, 1999).

### **Surgimiento de la captura de movimiento sin marcadores.**

La captura de movimiento más tradicional (óptica, electromagnética y electromecánica) se ha desarrollado y estudiado a lo largo de las últimas décadas, pero cada vez están surgiendo más tecnologías que mejoran la calidad de la data capturada con ellas. La gran mayoría de estos avances tecnológicos estaban reservados para el desarrollo de películas con grandes presupuestos, videojuegos y estudios de biomecánicas, debido a que en su mayoría necesitaban de entornos como laboratorios o sets complejos con equipos sofisticados, lo cual hacía el campo de la captura de movimiento exclusivo para los profesionales de la animación. Sin embargo, y a pesar de todo estos avances, una de las principales desventajas o fallas de estos sistemas es la necesidad de utilizar marcadores, lo cual aumenta el tiempo de preparación y configuración requerido para comenzar la captura de movimiento; además de molestar a los actores, quienes tienen que utilizar una abundancia de sensores o marcadores que en algunos casos pueden restringir su movimiento.

In addition, obtrusive sensors must be attached to actors and calibrated within the capturing system, resulting in limited and unnatural motion. In recent year the rise of computer vision and interactive entertainment opened the gate for a different type of motion capture which focuses on producing marker or mechanical sensorless motion capture.

(TRONG DUONG, 2012, p. iii)

De ahí que surja una de las tecnologías para lidiar con esta complicación, que es la captura de movimiento sin marcadores, la cual elimina una de las complicaciones de los sistemas de captura. Esta, como casi todas las otras tecnologías en este campo, surge en las universidades, entre las cuales está La Universidad De Stanford, cuyo programa sigue vigente, aunque sus avances se han enfocado en la utilización de la captura de movimiento sin marcadores en el área de la Biomecánica. Al punto que en la actualidad se puede decir que, la mayoría de la documentación académica sobre este tipo de sistema se deriva de la misma. Todos estos sistemas tienen sus raíces en estudios realizados por la NASA, institución que desarrolló las bases del rastreo óptico de movimiento sin marcadores, motivado a que tenía la necesidad de rastrear una diversidad de objetos, utilizando una variedad de cámaras y lentes de largo alcance, las cuales generaran imágenes de alta resolución del objeto rastreado, lo que le permitiría obtener más información y no solamente los datos de movimiento. Pero se debe reconocer que estos sistemas se logra desarrollar gracias al trabajo de George Veis, quien publicó en 1963 su

trabajo: *Optical tracking of artificial satellites*, el cual constituyó la base teórica para este tipo de sistemas (Veis, 1963).

Esta tecnología fue llevada por primera vez a la gran pantalla, con las películas de *The Matrix Reloaded* (Wachowski & Wachowski, 2003a) y *The Matrix Revolution* (Wachowski & Wachowski, 2003b), en este caso no se hizo captura de cuerpo entero, sino captura de movimiento facial, por eso algunos argumentan que es un campo diferente a la captura de movimiento de cuerpo entero, ya que requiere de software y equipos especiales. Como lo explica Borshukov (2005) en el trabajo que luego publicó del equipo de efectos especiales, allí explica la necesidad de realizar animación facial, pero sin utilizar los métodos tradicionales como deformaciones musculares, entre otras. Esto se debía especialmente a que tenían que recrear las caras de actores conocidos como: Keanu Reeves y Lawrence Fishburne. Para esto se enfocó en soluciones basadas en la aplicación de técnicas apoyadas en la imagen.

Our team had been very successful at applying image-based techniques for photorealistic film set/location rendering, so we decided to approach the problem from the image-based side again. We wanted to produce a 3-d recording of the real actor's performance and be able to play it back from different angles and under different lighting conditions. Just as we can extract geometry, texture, or light from images, we are now able to extract movement. Universal Capture combines two powerful computer vision techniques: optical flow and photogrammetry.

(Borshukov, Piponi, Larsen, Lewis, & Tempelaar-Lietz, 2005)

Este método de captura, que ellos denominan Universal Capture, se asemeja a la técnica que desarrollaron para crear el efecto de “*bullet time*”, en este método se busca rastrear el Flujo óptico de cada pixel grabado de manera 2D por el sistema de 7 cámaras, y los buscan interconectar basándose en la fotogrametría. Este método de captura requiere mucho poder de procesamiento y no es práctico para captura de movimiento de cuerpo completo. Pero a pesar de esto, si era capaz de generar imágenes foto-realista y lograr expresiones faciales muy naturales.

De manera similar a las otras tecnologías más antiguas para la captura de movimiento, los sistemas sin marcadores que se utilizaban eran sistemas hechos a la medida, para ser usados por profesionales en el campo de la animación y computarización. No existía algún programa o

sistema que se pudiera comprar e instalar de manera fácil y rápida, como se podía hacer con otros sistemas más estandarizados. Pero ya para la segunda mitad de la primera década del 2000, con el avance de las computadoras, las cuales eran cada vez capaces de procesar más data, y de manera más eficiente, comienzan a surgir las compañías que se especializaron en software para la captura de movimiento sin marcadores.



Fig. 1.19 a. Sistema de captura de movimiento de Organic Motion Openstage 2, capaz de captura de movimiento sin marcadores en tiempo real,

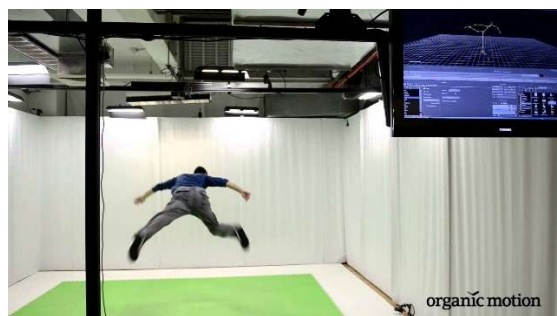


Fig. 1.19 b. Sistema Openstage 2 en uso, como se evidencia, el actor no tiene ningún tipo de ropa especial.

Esto no quiere decir que no existan programas profesionales, basados en la tecnología de captura de movimiento sin marcadores. Existen, uno de estos programas es el de la compañía *Organic Motion* (ver fig. 1.19), la cual ofreció una gama de productos para la captura de movimiento sin marcadores en el año 2007, esta compañía en la actualidad se ha enfocado en desarrollar la captura en tiempo real. Sus productos se han comercializado exitosamente, porque son más rápidos y económicos, al no necesitar marcadores.

### **Como funcionan la captura de movimiento sin marcadores.**

La idea de la captura de movimiento sin marcadores, se basa en los mismos principios de la triangulación que los sistemas ópticos con marcadores; es decir que buscan triangular la posición de un objetivo para poder así conocer su posición en el espacio y poder traducir esos movimientos del mundo real al mundo virtual. Para lograr esto, de manera similar a los otros sistemas ópticos, se pueden dividir en dos partes:

En la primera parte, el sistema de imagen óptica es responsable de convertir la luz en la imagen digital para que la computadora pueda procesar la data y convertirla en data de movimiento. Dependiendo del diseño puede variar desde algo tan simple como una cámara digital, hasta sensores de profundidad, basados en luz infrarroja que se utilizan principalmente

para aplicaciones sin marcadores. Las capacidades del sistema óptico determinan los límites del sistema de captura, ya que estos determinan el volumen de captura. La segunda parte es la computadora, donde se procesara los datos capturados por el sistema de imagen óptica, para convertirla en data de movimiento u otro tipo de información. Esto se logra al analizar la imagen, si es un sistema tradicional con marcadores busca identificar los marcadores, lo cual es tarea relativamente fácil, ya que estos fueron diseñados para que resalten del fondo, ya una vez identificados los marcadores estos se rastrean a lo largo del video. Pero en el caso de no tener marcadores que resalten del fondo, se necesita algo más complejo que los algoritmos, los desarrollados por Veis y sus derivados, sino que se entra al campo de la visión artificial. La visión artificial no es solamente el poder rastrear una bola de luz en un volumen de captura, porque un sistema óptico sin marcadores necesita para poder funcionar alguna forma, aunque sea rudimentaria, necesita saber qué es lo que ve. Según Klette la visión artificial es que la computadora a través de cámaras sea capaz de analizar y/o entender escenas en el mundo real.

In computer vision we may want to know how far away a building is to a camera, whether a vehicle drives in the middle of its lane, how many people are in a scene, or we even want to recognize a particular person—all to be answered based on recorded images or videos.

(Klette, 2014, p. vii)

Para poder lograr todo esto se tienen que desarrollar un software, que simule la capacidad humana de ver, pero no solo ver, porque ya hay muchas cámaras que son capaces de ver cosas que nosotros como humanos somos físicamente incapaces de ver, sin la ayuda de estos dispositivos. Es ir más allá de ver, es entender lo que se ve, y para eso se necesita un grado cada vez mayor de inteligencia computarizada, o más bien una computadora con simulación de una inteligencia, que está conformada por una gran cantidad de algoritmos, que permiten que la maquina pueda reconocer y hacer lo que se le ha enseñado. En los últimos años, la tecnología ha avanzado en el desarrollo de computadoras más poderosas, permitiendo que los algoritmos, especializados que son la base de la visión artificial se hagan más complejos, pero ampliando sus capacidades de análisis y mejoras en los sensores de las cámaras digitales, lo cual hace que estos sistemas tengan acceso a una cantidad cada vez mayor de data, mejorando así la precisión de los mismos. Esto ha dado la oportunidad para que la tecnología de visión artificial saliera de

los laboratorios y se aplicara en otros campos, diferentes a los de e la investigación y el puro desarrollo tecnológico.

There are significant contributions for the movie industry (e.g. the use of avatars or the creation of virtual worlds based on recorded images, the enhancement of historic video data, or high-quality presentations of movies)

(Klette, 2014, p. vii)

En el caso de la captura de movimiento sin marcadores, la visión artificial o computarizada, debe ser capaz de reconocer la silueta humana, y rastrearla a lo largo de los cuadros, además de esto también tiene que ser capaz de segmentar la silueta, de manera similar a cuando se segmenta con marcadores el cuerpo del actor, esta segmentación se basa en un esqueleto digital, que es una forma rudimentaria del humano. De esta forma, cuando un actor interviene en el set de captura de movimiento, la computadora será capaz de rastrear la silueta del actor y su movimiento en el espacio, a fin de luego traducir el movimiento del mundo real a data de movimiento que se puede aplicar a personajes en un mundo virtual. Esto no significa que haya ausencia de limitaciones para este tipo de sistemas, ya que dependiendo de las capacidades de procesamiento o de transferencia de data, un sistema puede que sea incapaz de realizar este proceso en tiempo real, como ocurre con algunos sistemas de bajo costo, otros sacrifican la precisión extra que les traería en no procesar en tiempo real, por la habilidad de procesar estos datos en tiempo real. Pero hay casos donde no es necesario un alto nivel de precisión, como cuando se utiliza la captura de movimiento sin marcadores para reconocer poses o gestos generales.

### **El uso masivo de la captura de movimiento sin marcadores.**

En la actualidad, cada vez más tecnologías altamente avanzadas se vuelven accesibles al ciudadano común, y en muchos casos, estos avances son más perceptibles que otros, como es el caso de los Smartphone, que cambiaron la sociedad entera, mientras que otros solo facilitaban cosas que ya hacíamos. Este hacer de la sociedad moderna, donde la tecnología es cada vez más accesible, no solo en bases económicas, sino en cuanto al conocimiento necesario para poder operar estas tecnologías, es lo que trajo la posibilidad de hacer captura de movimiento a las masas, pero esto no fue un proceso instantáneo, sino que ha sido un proceso de más de una

década. Una de las industrias que ha impulsado esta penetración de la captura de movimiento en la vida cotidiana, es la industria de videojuegos, aunque nunca fue bajo ese fin de que sus usuarios animaran sus propias producciones, sino para que estos pudieran interactuar con los juegos de manera más natural.

Entre las primeras insistencias esta la webcams que desarrollo Sony , tanto para su respectiva consola de la sexta generación, como Playstation 2, estos eran sistemas rudimentarios, incapaces de la captura de movimiento tridimensional, y no fueron muy bien acogidas por el público en general, pero ello obedeció más a que los juegos que utilizaban la tecnología eran pocos atractivos. También se debe tener en cuenta que durante los primeros años del siglo XXI, las tecnologías de la animación se hicieron más accesibles y aumentaron los programas de animación disponibles para los fanáticos de esa tecnología. Comenzaron a surgir comunidades en foros virtuales que empiezan a crear software libre para la realización de captura de movimiento. Además, poco a poco hay más pequeños estudios que se están dedicando a la animación y la producción de video juegos. Lo cual hace lógico el surgimiento de varias compañías que buscan solucionar sus necesidades. Entre estas hubo de diversos tipos, están las que buscaban satisfacer las necesidades de profesionales, y otras que se enfocaban más en el mercado amateur o artistas 3D independientes. Entre los primeros sistemas sin marcadores que salieron al mercado fueron el sistema de iPiSoft, el cual utilizaba de 3 a 6 cámaras en su primer iteración, como se mencionó antes funciona de manera similar a los otros sistemas ópticos; más tarde, pero en esos mismos primeros años del siglo XXI, Microsoft introduce al mercado el Kinect, que busca traer la captura de movimiento al hogar como una forma de interactuar con el mundo virtual de un video juego, sin la necesidad de un control o mando (ver fig. 1.20) . Por esto, cuando Microsoft presentó el Xbox Kinect en el 2010, un equipo diseñado para que sea capaz de captura movimiento en tiempo real y que funciona como un dispositivo *plug-n-play* y sin necesidad de marcadores. Este fue rápidamente acogido por los desarrolladores de software y sistemas, quienes necesitaban un dispositivo de bajo costo y fácil uso. Del Xbox Kinect se hará referencia con mayor profundidad en el próximo capítulo.



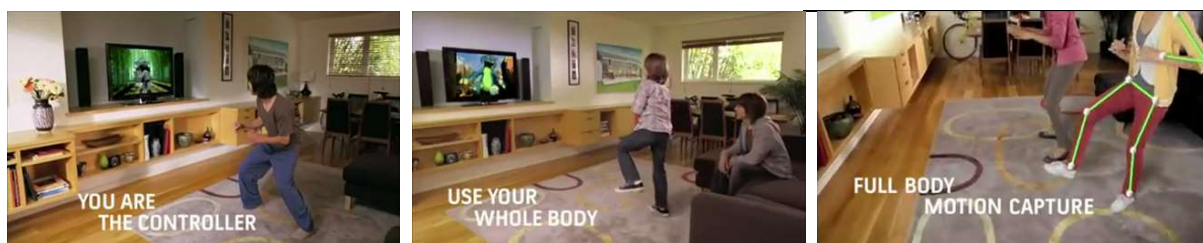


Fig. 1.20 Exerptos de un comercial del para el Xbox Kinect, donde se promocionan las habilidades de captura de movimiento del sistema, y las posibilidades de entretenimiento, sin la necesidad de un control o mando que este conectado a la consola.

Hoy en día, se puede decir que este sistema está en su infancia en comparación con otros sistemas que ya tienen años en el mercado; principalmente, porque los sistemas actuales se ven limitados por el procesamiento de data. El desarrollo de esta tecnología no se puede utilizar en aplicaciones en tiempo real, ya que no existe la capacidad de procesamiento lo suficientemente rápida para hacerlo, excepto que uno utilice una supercomputadora. Pero según la ley de Moore, mientras más años pasen, los procesadores serán más rápidos y se llegara a un punto que será, no solamente posible, sino que se podrá trabajar en una computadora estándar, de la misma forma que hace 30 años se necesitaba una computadora especializada para correr los programas de animación y captura de movimiento, la cual se podía conseguir solamente en ciertas universidades o compañías especializadas. En la actualidad, casi cualquier computadora en el mercado es capaz de correr el software de los sistemas ópticos con marcadores más avanzados, aunque tiene varias desventajas con los sistemas más sofisticados del mercado. Sin embargo, si se compara con los sistemas que están alrededor del mismo rango de precio, tienden a rendir de igual o mejor forma que los otros. Debido a su fácil acceso, este tipo de sistemas se han vuelto entre los más populares con pequeños estudios y productores independientes.

#### 1.5.1. 2.- Sistemas No-Ópticos:

Los sistemas no ópticos de captura de movimiento son mucho más variados que los sistemas ópticos, algunos de ellos funcionan bajo los mismos principios que los ópticos, mientras que otros sistemas no ópticos utilizan una variedad de sensores como potenciómetros y acelerómetros para capturar el movimiento. En este trabajo, se consideraron solo los Sistemas de Rastreadores Electromagnéticos y los Trajes de captura de movimiento electromecánico, porque son los que tienen más trascendencia y permanencia en la industria.

#### 1.5.1.2.3.- Sistemas de Rastreadores Electromagnéticos

Autores como Kitagawa y Windsor, al igual que Menache plantean que los sistemas de captura de movimiento electromagnéticos permiten tener seis grados de libertad. Los sistemas de medición electromagnéticos consisten en una serie de receptores que miden su relación espacial con un transmisor de cerca. Estos receptores o sensores se colocan en el cuerpo, están conectados a una unidad de control electrónica y en la mayoría de los casos por cables individuales.

En concreto, un rastreador magnético típico consiste en un transmisor, con 11-18 sensores, una unidad de control electrónico, y el software. Por su parte, la técnica de seguimiento magnética puede tener hasta 90 sensores y es capaz de capturar hasta 144 muestras por segundo. Su costo oscila entre \$ 5000 a \$ 150.000, lo que resulta considerablemente menos oneroso que los sistemas ópticos. El transmisor genera un campo electromagnético de baja frecuencia que es detectada por los receptores, tiene de entrada una unidad de control electrónica, donde se filtra y se amplifica para enviarla de inmediato a un ordenador central, allí el software resuelve la posición de cada sensor en x, y, y z coordenadas cartesianas y la orientación (guiñada, cabeceo y balanceo).

Es importante señalar que, rastreadores magnéticos como *Flock-of-Bird* (Multitud de pájaros) producidos por *Ascension Technology Corporation* usan (DC) campos electromagnéticos directos actuales, mientras que otros, como el *Polhemus Fastrak*, usa campos (AC) de corriente alterna. Estas tecnologías tienen problemas diferentes asociados con la conductividad metálica. Rastreadores de CA son muy sensibles al aluminio, cobre y acero al carbono, pero no es tan sensible a los de acero inoxidable o de hierro, mientras que los seguidores de (CC) corriente continua tienen problemas con los metales ferrosos, tales como el hierro, pero no con aluminio y cobre (Menache, 2010).

Muchos de estos problemas de conductividad son causados por la inducción de una corriente en el metal que crea un nuevo campo electromagnético que interfiere con el campo original emitida por el marcador. Estos nuevos campos se denominan corrientes de Foucault. Algunos rastreadores magnéticos utilizan algoritmos especiales para compensar estas distorsiones mediante la asignación de la zona de captura, pero estas calibraciones funcionan

sólo con áreas de problemas estáticos, predefinidos tales como las estructuras metálicas en edificios. En la mayoría de los casos, es mejor evitar cualquier metal de alta conductividad cerca de la zona donde se va realizar la captura de movimiento. Esta limitación hace que el rastreador magnético sea difícil de transportar a diferentes sets.

Algunas veces, se ha utilizado como parte de los dispositivos de *Puppeteering*, (títeres), donde un actor realiza los movimientos del cuerpo de un personaje con el rastreador magnético, mientras que otra persona realiza las expresiones faciales y la sincronización de labios, utilizando un rastreador de cara o un guante de datos. Al mismo tiempo, un titiritero puede animar los ojos del personaje, usando un simple ratón. También se utilizan para obtener posicionamiento global en otros sistemas, tales como trajes de inercia de captura de movimiento. En la tabla 1.2 se presentan las ventajas y desventajas del sistema de Rastreadores Electromagnéticos.

Tabla 1.2 Ventajas y desventajas de Rastreadores Electromagnéticos:

Ventajas	Desventajas
<ul style="list-style-type: none"> <li>• Salida de datos en tiempo real puede proporcionar una respuesta inmediata.</li> <li>• Posición y datos de orientación está disponible sin post-processing.</li> <li>• Rastreadores magnéticos son menos costosos que los sistemas ópticos, con un costo de entre \$ 5000 y \$ 150.000.</li> <li>• Los sensores no están ocluidos.</li> <li>• Es posible capturar múltiples actores que interactúan simultáneamente con múltiples configuraciones.</li> </ul>	<ul style="list-style-type: none"> <li>• La sensibilidad del rastreador de metal puede resultar en la salida irregular.</li> <li>• Intérpretes están limitados por cables en algunos casos.</li> <li>• Rastreadores magnéticos tienen una frecuencia de muestreo más baja que algunos sistemas ópticos o LPS.</li> <li>• El área de captura es más pequeño que es posible con otros sistemas.</li> <li>• Es difícil cambiar las configuraciones de marcadores.</li> </ul>

Fuente:(Menache, 2010) (Kitagawa & Windsor, 2008)

#### 1.5.1.2.4.- Trajes de captura de movimiento electromecánico.

Según, Menache los trajes Electromecánicos de captura de movimiento han estado presentes desde hace tiempo. Son sistemas basados en un grupo de estructuras unidas por potenciómetros, giroscopios, o dispositivos de medición angular similares, ubicados en las principales articulaciones humanas. Las versiones más recientes utilizan MEMS (Sistemas micro electromecánicos) sensores inerciales, colocados sobre una lycra o traje de spandex. La idea es que el traje mide todas las rotaciones de las extremidades humanas. Los potenciómetros son componentes que se han utilizado durante muchos años en la industria electrónica, en

aplicaciones tales como controles de volumen en radios de edad. Una corredera se mueve a lo largo de un elemento de resistencia en el potenciómetro para registrar una lectura potencial de voltaje variable, dependiendo qué porcentaje de la resistencia total se aplica a la tensión de entrada. Los sensores inerciales (MEMS) son también dispositivos de medición angular, pero son mucho más pequeños y más precisos. Son básicamente circuitos integrados, combinados con partes móviles micro-mecanizados.

Buenos ejemplos de dispositivos que utilizan la tecnología MEMS son el dispositivo de entrada de Nintendo Wii y el iPhone. Un gran inconveniente de los trajes de captura de movimiento es su incapacidad para medir traducciones globales. En la mayoría de los casos, se añade un sensor electromagnético o ultrasonido para la configuración de resolver este problema, pero se somete a la configuración de las mismas desventajas que los sistemas electromagnéticos o ultrasónicos, como la sensibilidad a los metales cercanos o mala precisión y deriva. Además, el diseño de estos dispositivos parte de la suposición de que la mayoría de los huesos humanos están conectados por articulaciones de bisagra simples, por lo que no tienen en cuenta las rotaciones no estándar. Las cuales son comunes a las articulaciones humanas, como en el hombro y antebrazo. Por supuesto, en realidad esto puede ser una ventaja si la configuración mecánica de un carácter digital en particular llama para tales tipos de restricciones. En la tabla 1.3 se presentan las ventajas y desventajas de los trajes de captura de movimiento electromecánico:

Tabla 1.3 Ventajas y Desventajas de Trajes de captura de movimiento electromecánico:

Ventajas	Desventajas
<ul style="list-style-type: none"> <li>• El rango de captura puede ser muy grande.</li> <li>• Trajes electromecánicos son menos costosos que los sistemas ópticos y magnéticos.</li> <li>• La demanda es portátil.</li> <li>• Recogida de datos en tiempo real es posible.</li> <li>• Los datos son de bajo costo para capturar.</li> <li>• Los sensores no pueden tener oclusión.</li> <li>• Es posible capturar múltiples artistas simultáneamente con múltiples configuraciones.</li> </ul>	<ul style="list-style-type: none"> <li>• Los sistemas tienen una frecuencia de muestreo baja.</li> <li>• Pueden ser molesto debido a la cantidad de hardware.</li> <li>• Los sistemas prestan restricciones en las articulaciones humanas.</li> <li>• La configuración de los sensores es fijo.</li> <li>• La mayoría de los sistemas no calculan traducciones globales.</li> </ul>

Fuente: (Menache, 2010)

Existen otros sistemas de captura de movimiento no ópticos, pero no se abordaron porque van más allá de la temática de la presente investigación, ya que ésta trata de las tecnologías de video juego como sistema óptico sin marcadores para la captura de movimiento.



## Capítulo II: Las tecnologías de los videojuegos como Sistema óptico de Captura de movimientos.

La captura de movimiento, la centrada especialmente en el movimiento humano, se ha realizado con una variedad de tecnologías a lo largo de las décadas, como se expuso en el capítulo anterior. En la actualidad surgen cada vez más avances, los cuales buscan reducir los costos de dicha captura, entre esos avances está la captura de movimiento sin marcadores. Una de las áreas pioneras donde se ubicó esta innovación a bajo costo fue en la industria de los videojuegos. La misma ha sido una de las grandes promotoras del uso de diversas tecnologías con propósitos de entretenimiento, desde sus inicios trajo diferentes innovaciones desde la computación a los televisores de los hogares alrededor del mundo.

A principios de este siglo, la industria de los videojuegos buscó lograr que la interactividad del jugador fuera más allá del control de la consola, para ello comenzaron a desarrollarse una variedad de accesorios, desde las alfombras de bailes de juegos como “*Dance Dance Revolution*”, hasta las primeras incursiones en la captura de movimiento como los sistemas de cámaras para el “*Playstation*”, que llevaron la captura de movimiento sin marcadores a los hogares, a través del “*EyeToy*” en el 2004, el cual rastreaba la silueta bidimensional del jugador para que pudiera interactuar con el juego. Pero no fue sino hasta la llegada del “*Nintendo Wii*”, a finales del 2006 que la idea de “motion-controllers” o controles a base de movimiento que comenzaron a cobrar fuerza, ya que el *Wii* presentó la idea de que la interactividad debía ser con todo el cuerpo como sus comerciales promovían. El *Wii* pasó a ser la consola mejor vendida de esa generación, con más de cien millones de unidades; esto hizo que, las otras compañías manufactureras de consolas buscaran desarrollar y mejorar la interactividad a través del “motion-control” con sus respectivos productos.

Los enfoques de las otras dos compañías rivales del *Wii* fueron diferentes, ya que una se basaba en tecnologías ya probadas, mientras que otra buscó ir más a lo no probado antes. Ambas anunciaron sus productos en el 2009, pero no fue sino hasta el 2010 que estos llegan al mercado. Sony logró desarrollar un sistema híbrido, donde mezcla un sistema óptico activo, en ese sistema el marcador está integrado a un mando o control con varios sensores, similares a los que se encuentran en los trajes de captura de movimiento con sensores (acelerómetros) y en

todos los Smartphones de alta gama, muy similar al control del *Wii*, aunque mucho más preciso; este control era el *Playstation Move*, el cual funcionaba en conjunto al *Playstation Eye*, creado en el 2007 como una webcam para el Playstation, sucesora del *EyeToy*. Esta tecnología se basa en cámaras que tienen sensores con una modalidad de color RGB, con la cual funcionan las mayorías de las cámaras a nivel mundial.

Por su parte, Microsoft desarrolló el *Xbox Kinect*, con el cual buscó romper el paradigma de que se necesitaba un control para poder interactuar con los videojuegos y que más bien (Xbox On, n.d.), como decía su campaña publicitaria “*you are the controller*” tú eres el control remoto (Xbox On, n.d.). Buscaron desarrollar un dispositivo sin igual para el alcance del consumidor común, un sensor de profundidad tridimensional, con la capacidad de captar el movimiento del usuario en tiempo real, con la precisión suficiente para que el jugador pudiera interactuar con el mundo virtual que ve en la pantalla de su televisor. Esta tecnología va más allá de las capacidades de las típicas cámaras RGB, debido a que estas funcionan con un sensor infrarrojo y crea una nueva forma de capturar data, llaman RGB-D, ya que es capaz de grabar en tres dimensiones (Regazzoni, de Vecchi, & Rizzi, 2014).

En este trabajo se abordaron las propuestas de estas dos compañías, pero con énfasis en el Kinect de Microsoft, porque constituye lo más novedoso en el área.

## **2.1.- Accesorios de videojuegos, como sistemas de captura de movimiento.**

En este punto, se consideraron como accesorio de video juegos los siguientes: el Playstation Eye y Move, El Kinect y Múltiples cámaras (RGB) o múltiples Sensores de Profundidad vs (RGB-D), el cuál es el más viable para la captura (Regazzoni et al., 2014).

### **2.1.1. El Playstation Eye y Move.**

Sony, lanzó al mercado el *Playstation Eye*, como sucesor del *Eye Toy*, la webcam de su previa consola. La cámara, como se dijo anteriormente, es comparable a una webcam de computadora. El *PlayStation Eye* es capaz de capturar con velocidades de cuadro de 60 Hz con una resolución de píxel de  $640 \times 480$ , y 120 hercios a  $320 \times 240$  píxeles, que es "cuatro veces la resolución" y "dos veces el marcos tasa" del *EyeToy*, según Sony. Del mismo modo tiene el doble de la sensibilidad a la luz que su predecesor, gracias a la colaboración de Sony con *OmniVision Technologies* que se diseñó un sensor con píxeles de mayor tamaño, lo que



permitió un funcionamiento bajo una luz más eficaz. La cámara cuenta con un objetivo zoom de enfoque fijo, ajustable con dos configuraciones seleccionadas manualmente, girando el tubo de la lente; la cámara *PlayStation Eye* puede establecerse en un campo de visión (punto rojo) similar a la de la EyeToy, a  $56^\circ$  para el primer plano de encuadre en las aplicaciones de chat, o en un campo de  $75^\circ$  vista (punto azul) por mucho tiempo con encuadre de tiro en aplicaciones de juegos interactivos físico (“Sony Eye specs,” n.d.) (“Sony PlayStation Eye Camera Specs,” n.d.). El *PlayStation Eye* es capaz de dar salida de vídeo sin comprimir, lo que lo convirtió en un candidato ideal para aplicaciones, como la captura de movimiento sin marcadores. Más adelante se explicara este aspecto. La referida cámara tuvo un sucesor, que es compatible con el Playstation 4, pero no tuvo la proyección para la captura de movimiento de la misma forma que su predecesora, debido en parte a su costo más elevado y a que no posee drivers para que ésta sea utilizada con la computadora (“PLAYSTATION EYE - Q+A | Three Speech,” n.d.).

Siguiendo con lo explícito por varios periodistas y analistas (Saltzman, 2010) (Kuchera, 2010) Sony, debido a presión de la competencia decide aventurarse a la idea de *motion controllers*, saca al mercado el PlayStation Move, al igual que el control inalámbrico estándar del PlayStation 3 (Sixaxis, DualShock 3), el PlayStation Move usa un módulo de radio comunicación inalámbrica de Bluetooth 2.0 y una batería interna de iones de litio que se carga a través de un puerto USB Mini-B. El PlayStation Move cuenta con una variedad de sensores internos para medir la orientación, además cuenta con una esfera de goma en la parte superior que puede brillar en cualquier color de la amplia gama RGB, ya que utiliza diodos emisores de luz (LEDs), sirve como un marcador activo, la posición de la cual se puede seguir a lo largo del plano de la imagen por la cámara. La forma esférica uniforme y de un tamaño conocido de la luz, permite que el sistema puede determinar la distancia del control o mando de la cámara, simplemente calculando el tamaño de la imagen de la luz, lo que permite que la posición del controlador pueda ser rastreado en tres dimensiones con una alta precisión y exactitud (Marks, 2001)

El Playstation Move, en su interior cuenta con una gama variada de sensores. Posee un par de sensores inerciales, un sensor acelerómetro lineal de tres ejes y un sensor de velocidad angular de tres ejes, que se utilizan para realizar un seguimiento de la rotación, así como del movimiento global del mismo. Tiene también un magnetómetro interno que se utiliza para

calibrar la orientación del control o mando contra el campo magnético de la Tierra y ayudar a corregir contra el error acumulativo (derivado) por los sensores inerciales. Además, posee un sensor de temperatura interna que se utiliza para ajustar las lecturas de los sensores inerciales contra los efectos de temperatura. Los sensores inerciales pueden ser utilizados para la navegación en los casos que el seguimiento de la cámara sea insuficiente, como por ejemplo cuando el control o mando está oculto detrás de la espalda del actor, o en algún momento de oclusión.

### 2.1.2.- El Kinect para la captura de movimiento

El Xbox Kinect es un sistema óptico de captura de movimiento sin marcadores, acogido por la industria de videojuegos, los métodos más tradicionales de captura de movimiento que requieren marcadores para sistemas ópticos son utilizados por la industria de los videojuegos en el área de la producción, pero en el pasado en el ámbito de la interacción entre jugador-juego, ya se había intentado también. Sin embargo, no buscó rastrear el movimiento del cuerpo entero del jugador. Esta necesidad llevó a innovaciones en el campo de la captura de movimiento sin marcadores, ya que es el único sistema en el que el usuario, bien sea como un jugador o un actor no necesita ningún tipo de instrumento en su cuerpo, lo cual facilita el uso del mismo. El propósito principal de esta industria, como ya se mencionó, era lograr traducir los movimientos del jugador con acciones y movimientos en el videojuego, sin que se necesitara algún tipo de equipo extra, como controles o reflectores de los sistemas más tradicionales, aquí se requiere solamente de la cámara (Kinect) y equipo (Xbox360) donde va a estar conectado. (Ver fig. 2.1)



Fig. 2.1 a. Kinect para el Xbox 360



Fig. 2.1 b. Combo Xbox 360 Slim con Kinect y juego incluido.

Este sistema fue logrado por Microsoft, y fue el primer aparato de su tipo, diseñado para la captura de movimiento, destinado al mercado masivo mundial de los videos juegos. Según Esler (2012) el Xbox Kinect fue creado específicamente como producto para el mercado

norteamericano y destinado a la consola de video juegos Xbox 360, en noviembre del 2010, y según el mismo autor, en febrero del 2012 se ofrece para computadoras que corran el sistema operativo Microsoft Windows 7 en adelante. El sistema óptico de captura de movimiento Xbox Kinect, según (Khoshelham & Elberink, 2012) permite a los usuarios controlar e interactuar con la consola, sin necesidad de tener contacto físico con controles de videojuegos tradicionales, lo cual se hace mediante una interfaz natural para el usuario que reconoce gestos, comandos de voz, u objetos e imágenes.

The Kinect is a depth camera. Normal cameras collect the light that bounces off of the objects in front of them. They turn this light into an image that resembles what we see with our own eyes. The Kinect, on the other hand, records the distance of the objects that are placed in front of it. It uses infrared light to create an image (a depth image) that captures not what the objects look like, but where they are in space.

(Borenstein, 2012)

Lo impresionante del Kinect no es que pueda hacer captura de movimiento, sino que redujo la cantidad de equipos necesarios para hacerla, aunque su fin fue específicamente la comodidad del usuario, ya que como industria del entretenimiento buscaba un sistema simple de calibrar y de fácil operación, su mercado no era el profesional, sino el mercado del entretenimiento, cuyos usuarios están acostumbrados a productos que se enchufen y se utilicen de inmediato; este era su punto clave, porque si el equipo era muy complicado de usar, la gente no lo compraría. Se puede decir entonces que, la principal razón del surgimiento de la tecnología de captura de movimiento sin marcadores en el ámbito de los videojuegos, al contrario de la mayoría de los otros avances que surgieron relacionados al campo de la medicina, defensa, el cine y televisión, tuvo necesidades totalmente diferentes.

El Kinect fue desarrollado principalmente por Microsoft para permitir la interacción entre usuario/consola de video juego, sin la necesidad de usar un control. Buscando utilizar un sistema más natural a través de gestos, movimientos corporales y la voz del jugador. Este sistema se basa en una tecnología de software desarrollada por RARE, que era una subsidiaria de Microsoft y con la tecnología de *range imaging camera*, desarrollada por la compañía israelí, llamada “PrimeSense”, la cual desarrolló un sistema capaz de interpretar gestos a través de una cámara infrarroja con un microchip especial que le permite rastrear los movimientos de objetos individuales en tres dimensiones.

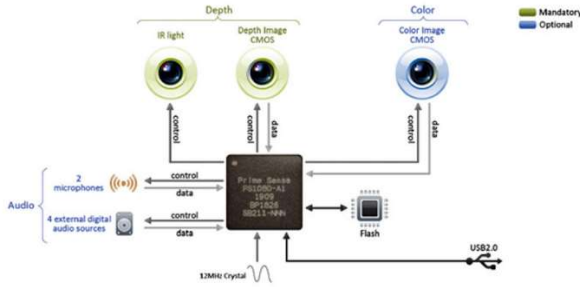


Fig. 2.2 Diagrama de los componentes que componen al Kinect para el Xbox 360.



Fig. 2.3 Kinect sin la carcasa, donde se pueden apreciar los componentes de izquierda a derecha, Sensor CMOS a color, Sensor de profundidad CMOS monocromatico y luz infraroja.

Físicamente el sensor Kinect es una barra horizontal conectada a una base pequeña, con un pivote motorizado y está diseñado para ser colocado longitudinalmente por encima o por debajo del televisor. El dispositivo cuenta con una cámara RGB (ver fig. 2.2), un sensor de profundidad (infrarrojo), y un conjunto de micrófonos. Gracias a los cuales es capaz de la captura de movimiento de cuerpo completo en un espacio tridimensional, con reconocimiento facial y capacidades de reconocimiento de voz. El sensor de profundidad se compone de un proyector de láser infrarrojo, combinado con un sensor CMOS monocromáticas, que captura los datos de vídeo en 3D bajo cualquier condición de luz ambiente (ver fig. 2.3). El rango de detección del sensor de profundidad es ajustable, y el software de Kinect es capaz de calibrar automáticamente el sensor basado en el juego y el entorno físico del jugador, además con capacidad para detectar la presencia de muebles u otros obstáculos.

En su aspecto más técnico el Kinect, de acuerdo a lo descrito presenta las siguientes características (Melgar & Diez, 2012):

- Las salidas de los sensores de video del Kinect tienen una velocidad que varía entre 9 Hz a 30 Hz, dependiendo de la resolución.
- El flujo de vídeo RGB predeterminado utiliza una resolución VGA de 8 bits (640 × 480 píxeles) con un filtro de color Bayer, pero el hardware es capaz de resoluciones de hasta 1280x1024 (a una velocidad de fotogramas más baja) y otros formatos de color como UYVY.
- El flujo de vídeo sensor de profundidad, monocromático es en resolución VGA (640 × 480 píxeles) con una profundidad de 11 bits, que ofrece 2.048 niveles de sensibilidad.

- El Kinect puede también transmitir la vista desde su cámara de infrarrojos directamente (es decir: antes de que se haya convertido en un mapa de profundidad) como 640x480 vídeo o 1280x1024 a una velocidad inferior.
- El sensor Kinect tiene un límite práctico, que va de 1.2 a 3.5 m (3.9 a 11.5 pies) de distancia cuando se utiliza con el software de Xbox. El área requerida para jugar Kinect es más o menos 6 m<sup>2</sup>, aunque el sensor puede mantener el seguimiento a través de una amplia gama de aproximadamente 0,7-6 m (2.3-19.7 pies). El sensor tiene un campo de visión angular de 57 ° en horizontal y 43 ° en vertical, mientras que el pivote motorizado es capaz de inclinar el sensor hasta 27°, ya sea hacia arriba o hacia abajo. Por consiguiente, el campo horizontal del sensor Kinect en la distancia de visualización mínima de ~ 0,8 m (2,6 pies) es ~ 87 cm (34 in), y el campo vertical es ~ 63 cm (25 pulgadas), resultando en una resolución de poco más de 1,3 mm (0,051 in) por píxel

La compañía Microsoft tres años después de sacar al mercado el Kinect original, logra sacar una versión actualizada con mayor capacidad de procesamiento de datos y mejor resolución para su más reciente consola de videojuego, este es el Xbox One. El Kinect para el Xbox One cuenta con una serie de cambios en comparación con la primera generación (Xbox 360). El nuevo Kinect ya no contiene un motor para el ajuste del ángulo automático, y debe ser ajustado manualmente. Al igual que con el Kinect original, el sensor utiliza infrarrojos para leer su entorno, pero tiene una mayor precisión respecto a su predecesor, gracias a que puede procesar 2 gigabits de datos por segundo. El dispositivo cuenta con una cámara de 512x424 pixel del tipo time-of-flight; que es un tipo de radar visual, el nuevo sensor tiene un mayor campo de visión, reduce la distancia necesaria entre el jugador y el sensor para un rendimiento óptimo(Khoshelham & Elberink, 2012). En comparación con los 2,4 m mínimo de Kinect para Xbox 360, Microsoft recomienda una distancia mínima de 1,37 m del Kinect para Xbox One.

El nuevo Kinect también incluye una cámara de vídeo 1080p de resolución que se puede utilizar para la grabación de vídeo y charla, en comparación con la de resolución VGA de la versión de Xbox 360. Se puede rastrear de hasta 6 esqueletos (usuarios) a la vez, rastrear de gestos realizados con un control de Xbox; además utiliza una matriz de cuatro micrófonos que permite un más amplio soporte para comandos de voz.

Es importante señalar lo indicado por (Weinberger, Sep. 8, 73, & 23, n.d.) al referir que, el Kinect, a pesar de ofrecer al jugador una mayor libertad de movimiento e interactividad que los productos de sus competidores, y haber sido un éxito económico al lograr vender más de 24 millones de unidades para el año 2013, no termino revolucionando el mundo de los videojuegos como sus creadores esperaban; esto debido a que el jugador nato de videojuego prefiere métodos de interactividad tradicional, que no requieren de mucho movimiento físico. Sin embargo, desde su llegada al mercado masivo el Kinect fue acogido por la comunidad mundial de inventores e innovadores, especialmente en los departamentos de tecnología de varias universidades alrededor del mundo; allí con el pasar de los años se desarrollaron varios proyectos interesantes alrededor de la capacidad de captura de movimiento y sensor de profundidad (Walker, 2012). Entre ello se destacan, los siguientes:

-El grupo de programadores del Robot Locomotion Group de MIT, están desarrollando drivers, para que el Kinect se pueda utilizar como un Motion Controller para una interface de usuario similar a lo que se veían en la película *Minority Report* (MITCSAIL, 2010).

-Durante mayo del 2011, una tienda de Topshop en Moscú estableció un quiosco de Kinect que podía superponer una colección de vestidos sobre una grabación en vivo de los clientes, funcionando como un espejo virtual. A través del seguimiento automático, la posición y rotación de los vestidos virtuales, estos se actualizan incluso cuando los clientes se daban vuelta para ver la parte posterior del vestido (Sterling, 2011).

-Investigadores de la Universidad de Minnesota han utilizado el Kinect para medir una gama de síntomas del trastorno en los niños, la creación de nuevas formas de evaluación objetiva para detectar condiciones tales como: el autismo, el trastorno por déficit de atención y trastorno obsesivo-compulsivo (“University of Minnesota researchers studying the use of robots and computer vision to diagnose mental disorders in children,” 2011).

Muchos de estas aplicaciones e innovaciones que salieron en los primeros meses, tras la llegada al mercado del Kinect, se realizaron sin drivers oficiales, fue el 16 de junio del 2011 que sale el Software Developers Kit (SDK) para Windows 7, aunque no fue hasta el año 2012 que sacaron una versión de Kinect para computadora, este kit, de acuerdo a las notas de la versión, ver anexo 1, incluía los drivers para el Kinect, al igual que las herramientas necesarias

para la creación de aplicaciones con C++, C# o Visual Basic a través de Microsoft Visual Studio 2010 que utilizan algunas de las siguientes especialidades:

-Flujo de datos de los sensores en bruto: El acceso a los flujos de bajo nivel del sensor de profundidad, sensor de la cámara de color, y matriz de micrófono de cuatro elementos.

-Rastreo esquelético: La capacidad de realizar un seguimiento de la imagen esqueleto de una o dos personas que se desplazan dentro del campo de vista Kinect para aplicaciones a bases de gestos.

-Capacidades avanzadas de audio: las capacidades de procesamiento de audio incluyen supresión sofisticada de ruido acústico y cancelación de eco, formación de haz para identificar la fuente de sonido actual, y la integración con la API de reconocimiento de voz de Windows.

-Muestra de código y documentación.

De ahí que, para mayo del 2015 sacan una actualización que añade varias nuevas características al SDK del Kinect en la versión 1.5, estas son las siguientes, de acuerdo a lo reseñado en las notas de la versión que forman parte del anexo 1:

- Kinect para Windows 1.5 SDK incluye 'Kinect Studio' una nueva aplicación que permite a los creadores de software grabar, reproducir y depurar clips de los usuarios interactuando con aplicaciones.
- Implementa un nuevo sistema esquelético "sentado" o de "10 - articulaciones" que le permitirá a las aplicaciones rastrear: cabeza, cuello y brazos de un usuario del Kinect – aunque esté sentado o de pie; lo cual trabajaría tanto en el modo estándar y el modo cercano.
- Se añaden cuatro nuevos idiomas para el reconocimiento de voz - francés, español, italiano y japonés. Además, se añade soporte para dialectos regionales de estos idiomas junto con el inglés.
- Se hace disponible para la descarga en los siguientes países: Hong Kong, Corea del Sur y Taiwán en mayo y para Austria, Bélgica, Brasil, Dinamarca, Finlandia, India, Países

Bajos , Noruega, Portugal, Rusia , Arabia Saudita , Singapur , Sudáfrica , Suecia , Suiza y los Emiratos árabes Unidos en junio.

### 2.1.3.- Múltiples cámaras (RGB) o múltiples Sensores de Profundidad vs (RGB-D) ¿cuál es el más viable para la captura?

En la actualidad, como se ha visto, existen dos vertientes para la realización de captura de movimiento sin marcadores, la que utiliza cámaras convencionales, similares a las que se utilizan en los sistemas ópticos más tradicionales, simplemente cambia en como el software procesa la data capturada, por otro lado la captura a base de los sensores de profundidad, la más reciente de estas dos tecnologías. Por consiguiente, se debería analizar cuáles son las posibles ventajas de utilizar unos sistemas con múltiples cámaras o con múltiples sensores de profundidad. Regazzoni (en su trabajo *RGB cams vs RGB-D sensors: Low cost motion capture technologies performances and limitations*, buscó averiguar justamente eso, su objetivo fue:

...benchmark analysis that compares RGB and RGB-D technologies used to track performing people in a variety of conditions. In order to contrast the solutions, several different tasks have been selected, simultaneously captured and post-processed exactly in the same way. The test campaign has been designed to evaluate pros and cons according to the most important feature of a motion capture technology, such as volume of acquisition, accuracy of joint position and tracking of fast movements.

(Regazzoni et al., 2014)

Básicamente el objetivo de su trabajo era comparar a ambos tipos de recolección de data para ver cuál era el más preciso, siguiendo un rigor científico en el cual se evalúan en condiciones variadas las capturas hechas de manera simultánea y el post-procesado exactamente de la misma manera. Es oportuno señalar que tanto en el trabajo de Regazzoni como en el presente trabajo, solo se están tomando en cuenta las tecnologías de captura de movimiento de bajo costo, y cuya accesibilidad está disponible al público general, ya que hay soluciones disponibles para mercados profesionales, pero cuyo costo es una limitante para las pequeñas compañías o amateurs. Específicamente, en su trabajo Regazzoni, (2014) utilizó el software de IPiSoft, del cual se habla con más detalle en la próxima sección del capítulo. Entre los hallazgos de dicho trabajo están los presentados en la tabla 2.1



TABLA 2.1: Comparación entre Xbox Kinect RGB-D y Playstation Eye (RGB)

	Xbox Kinect RGB-D	Playstation Eye (RGB)
Espacio	Requiere un espacio menor, un mínimo de 9m <sup>2</sup>	Requiere de más espacio
Volumen de captura	El volumen se ve limitado a 1 á	Es capaz de tener una mayor área de captura
Precisión	No hay ninguna ventaja entre ambos sistemas en cuestión de la precisión de la data capturada.	
Calibración	Su calibración es mucho más sencilla y rápida.	Requiere una calibración complicada y prolongada
Facilidad de instalación	Ya que solo se requieren instalar 2 sensores, es más sencilla	Es más compleja, ya que generalmente implica el uso de extensiones USB especializada y un total de 6 cámaras a instalar.
Facilidad de uso	No requiere iluminación especial o uniforme, y debido a su fácil instalación y calibración es el más fácil de utilizar de los 2.	Requiere iluminación uniforme. Pero al tener más cámaras, es más fácil evitar alguna posible oclusión.

Fuente: (Regazzoni et al., 2014)

Como podemos ver el Kinect, es el más ventajoso, esto se debe principalmente a que es más fácil de utilizar, lo cual se debe en parte, a que se utiliza de la manera para la cual fue diseñado, aunque no sea bajo el mismo fin; Microsoft diseñó el Kinect como un dispositivo capaz de realizar captura de movimiento lo suficientemente preciso para que pueda reconocer los movimientos de 1 o 2 jugadores, que estos puedan interactuar con el juego y que sea de fácil uso. Porque la intención de Microsoft fue lograr que el usuario sea capaz con solamente conectarlo y jugar de manera inmediata. Por su parte la cámara del Playstation Eye, aunque es capaz de capturar movimiento de una manera rudimentaria cuando es utilizada con la consola para la cual fue diseñada, fuera de este ámbito es utilizada para emular los sistemas profesionales, como los de Organic Motion. La razón por la cual es utilizada para la captura de movimiento por los interesados de la captura de movimiento o por IPiSoft, es que es una cámara de bajo costo, puede conseguirse con facilidad en cualquier lugar del mundo el mismo modelo, debido a que Sony es una gran compañía transnacional y lo ha hecho posible. No obstante ello, si se debe valorar que el Playstation Eye permite tener un volumen de captura mayor, en especial cuando la situación lo requiere. Para ello, se debe disponer de suficiente área para

instalar un estudio de captura permanente, de esta forma muchas de las desventajas del Playstation Eye desaparece. Igualmente, vale la pena mencionar que, solo uno de los programas de captura de movimiento no es compatible con las cámaras Playstation Eye, pues lo único que cambiaría con los otros programas sería el proceso de calibración

Es importante aclarar que, éste trabajo se enfocó en el uso del Kinect para la captura de movimiento. La razón de esto estuvo en que el Kinect por su cuenta ya es un sistema de captura de movimiento, porque el SDK de Microsoft tiene o da las herramientas necesarias para procesar la data en bruto capturada por el Kinect y pasarla a programas como Motion Builder de Autodesk y aplicarla a una animación. Sin embargo, resulta válido decir, que este no es un proceso fácil, requiere de un alto nivel de conocimiento técnico, no solo de programación, sino también de animación. Debido a ello, desde este punto en adelante, este trabajo se centró, en la captura de movimiento a través del Kinect.

## **2.2. Software con que funciona el Kinect.**

A pesar de que hay formas para crear software propio a través del SDK, este trabajo se enfocó en el software, ya establecido en el mercado de la captura de movimiento sin marcadores, se trata del software de la compañía IPISOFT.

### **2.2.1.- iPiSoft Captura de Movimiento para las Masas.**

La razón de la escogencia de este software obedece a que esta compañía desde sus inicios buscó integrar diferentes equipos para la captura de movimiento, de modo de hacerlo más accesible a los usuarios. En cambio las otras compañías buscaban que sus usuarios utilizaran equipos desarrollados por ellas mismas y que la creación de las instalaciones sean supervisadas por sus técnicos, bajo el fin de lograr mayor precisión en la captura de movimiento, pero sacrifican una mayor accesibilidad en el mercado, lo cual acarrea un costo muy elevado y que solo se puede utilizar en los países donde están esas compañías (Avlastimova, 2010).

#### **2.2.1.1.- Historia de la compañía iPiSoft**

La compañía de IPiSoft fue fundada por Mikhail Nikonov y Pavel Sorokin en el 2008, buscaba desarrollar una solución para la captura de movimiento que sea accesible para gente que se está iniciando en el mundo de la animación y para pequeños estudios. Aunque en sus inicios, esto no fue el caso, porque Nikonov como un graduado del Instituto de Física y

Tecnología de Moscú (MIPT por sus siglas en inglés), lo que quería era crear una compañía que hiciera software de negocios, programas pocos complicados y con usos más limitados, cuando se le compara con un programa para la captura de movimiento. Se pudiera decir que, esta falta de experiencia previa en el mundo de la animación, fue la que quizás les permitió tener una nueva perspectiva y conseguir soluciones poco convencionales, como es la captura de movimiento sin marcadores, porque para cuando la compañía había comenzado a trabajar en ello, en el 2008 no era un tema muy tratado en el mundo de la animación. Y fue esta posición de creadores ajenos al entorno de la animación, la que les permitió utilizar equipos que son más comunes alrededor del hogar, como las cámaras del estilo webcam, para realizar sus experimentos en lugar de las cámaras especializadas con las que trabajan las personas que ingresan al mundo de la producción audiovisual, de manera convencional y quienes ignorarían este tipo de cámaras, porque la considerarían poco útiles (Avlastimova, 2010).

Esta idea de desarrollar un programa que facilite la captura de movimiento, no surge hasta que Nikonov buscando crear un video juego como un pasa tiempo suyo, se dio cuenta que iba a necesitar animar personajes humanoides en 3 dimensiones. Para esto buscó cuáles sistemas estaban disponibles en el mercado para la captura de movimiento, se dio cuenta sus altos costos, con precios que comienzan por los diez mil dólares en adelante, y lo complicado que resulta operar los sistemas actuales de captura de movimiento. En este sentido, Nikonov, planteaba que *“Therefore, the idea to develop a system that would work on inexpensive equipment, and would be easy to operate.”* (Ilyin, 2010). Y precisamente en junio del 2010 iPiSoft sacó al mercado la primera versión de su programa, después de una etapa de beta abierta que duro varios meses, para este momento el Kinect todavía no había salido al mercado, y como desde sus inicios iPiSoft buscó hacer más accesible a las masas el uso de la captura de movimiento, integró diferentes equipos para lograr la captura de movimiento, de modo que sus usuarios tuvieran libertad a la hora de crear sus instalaciones y no dependieran de la supervisión de las compañías, por ello logró un sistema óptico de captura de movimiento más práctico. Su sistema se inició recomendando el uso de la cámara Playstation Eye, ya que era una cámara estándar que se encontraba en el mercado mundial y como eran utilizadas con la consola de Sony la Playstation 3. Esto permitía mantener un sentido de homogeneidad en el tipo de cámaras a utilizar, al igual que aseguraba que las mismas estuvieran disponibles en todo el mundo, dando así al inicio a

algo que distinguirá a iPiSoft del resto de las compañías en el ámbito de captura de movimiento, ya que desde entonces han basado siempre sus diseños en el uso de hardware, principalmente diseñados por otras compañías para ser utilizados en el mundo de los videojuegos, y de este modo hacerlo más accesible a la gente no especializada en el campo (Ilyin, 2010).

Siguiendo esta misión de llevar la captura de movimiento a las masas, fue que decidieron trabajar con los sistemas ópticos sin marcadores, no solo porque esto les permite el uso de cámaras digitales de bajo costo, sino porque también facilita el trabajo en la fase de pre captura, que es cuando se compara con los sistemas ópticos tradicionales, resultando más sencilla y rápida, ya que no necesitan trajes especiales, ni la colocación de marcadores, sean estos activos o pasivos (Ilyin, 2010). Este sistema se basa en no tener marcadores, éste rastrea cada pixel y lo utiliza como si fuera un marcador, de esta forma el sistema reduce las oportunidades que haya oclusión de alguna parte del cuerpo, reduciendo así la cantidad de cámaras necesarias (3 o 4) para operar el sistema. Mientras que si se les compara con sistemas que son ya estándares en la industria de la captura de movimiento, como los de Vicon, que utilizan 8 cámaras especializadas y hay sistemas que pueden llegar a tener cientos de ellas.

For comparison, existing systems use eight or more expensive special cameras, and the process of applying markers on the actor's very time consuming. In the traditional approach, even training courses for applying markers on the actor's worth more than our entire system.

(Avlastimova, 2010)

En sus inicios iPiSoft quería basar su programa alrededor de una sola cámara, específicamente una cámara del tipo estéreo para poder grabar video en 3 dimensiones, pero tal como refiere Nikonov (Avlastimova, 2010) en ese momento no estaba disponible una cámara con esas capacidades en el mercado. “The only thing that prevents us to release a single-camera version, is that conventional camera system works poorly, as appropriate three-dimensional camera is not on the market.” Nikonov (Ilyin, 2010); esto se debe a que las cámaras que para ese momento estaban disponibles en el mercado no tenían la capacidad para crear un “depth map” con la precisión necesaria a la distancia necesaria para emplearlas para la captura de movimiento, como dice Nikonov con respecto a las cámaras, según una entrevista que le hicieron en el 2010

"They're all set up, mainly for shooting from a distance of a meter or two, that is, that man shoot at close range. And we need a distance at least four or five meters from the camera to the actor."

Nikonov (Avlastimova, 2010)

El Kinect aunque salió al mercado a finales del 2010, no fue integrada al software de iPiSoft hasta el 6 de marzo del 2012, esto debido, como se dijo en la sección anterior, a que Microsoft no saca drivers oficiales para utilizar el Kinect en Windows hasta febrero del 2012. La compañía si había experimentado con el Kinect, utilizando los drivers no oficiales, pero estos nunca pasaron a ser oficiales y solo salió en forma de un beta, a fin de tener el sistema casi listo para cuando los drivers oficiales salieran al mercado y la implementación fuese lo más rápida posible. Es importante aclarar que el Kinect es un sensor de profundidad, capaz de hacer un "Depth Map", lo suficientemente preciso para realizar captura de movimiento con un solo aparato. Aunque, se recomienda utilizar más de uno, si los movimientos son complejos, y con la versión más actualizada, se pueden utilizar 4 equipos Kinect para realizar captura de movimiento con la mayor precisión posible.

Es importante precisar que el sistema de captura de movimiento de IPISOFT tiene ahora dos formas de realizar la captura, uno con cámaras Playstation Eye y el otro con el sensor Kinect, los cuales no se pueden utilizar combinados; es decir no se pueden utilizar varias cámaras Playstation Eye al lado de un sensor Kinect, ya que cada equipo se calibra de manera diferente, y el programa procesa los datos de cada uno de manera diferente. El usuario que disponga de uno de estos equipos puede usar el sistema IPISOFT. La razón es muy sencilla, como dice el lema de la compañía "Motion Capture for the masses", es un principio de la compañía tener un producto que facilita la experiencia del usuario y que no requiera conocimiento técnico extensivo para poder operarlo; de ahí que desde sus inicios se decidió no utilizar marcadores, ya que estos son incómodos para el actor y todo el proceso de aplicarlos es largo y tedioso, además de necesitar cierto nivel de capacitación para utilizarlo efectivamente.

En los últimos años, los usuarios le han solicitado a IPiSoft la posibilidad de hacer captura de movimiento con utilería, ya que la forma actual en que funciona la captura de movimiento sin marcadores busca reconocer el patrón de la figura humano y no la figura humana

más una pieza de utilería. IPiSoft para remediar esto volvieron al mundo de los videos juegos, a fin de ver que accesorio podría ayudarlo a resolver este problema, allí es cuando ven que Playstation para hacerle la competencia a Microsoft saca al mercado el Playstation Move, una especie de control que cuenta con un marcador activo, es decir una luz, que permite rastrear con mayor posibilidad a piezas de utilería y además posee una serie de sensores internos que transmite datos sobre su rotación, velocidad entre otros, lo cual se puede integrar al sistema y así permitir incorporar utilería a las sesiones de captura de movimiento. También es relevante mencionar que, cuando este producto salió al mercado aparecieron una serie de accesorios con los que hacían más fácil la inmersión del jugador al video juego, ya que estos se semejaban a espadas, pistolas y otras implementos.

#### 2.2.1.2. - Funcionamiento del software

El sistema de iPiSoft se califica principalmente como un sistema óptico sin marcadores, aunque tiene el potencial de ser un sistema híbrido, con la implementación de los controladores de movimiento que permiten la implementación de utilería. Este sistema consiste en dos programas diferentes, el primero es iPi Recorder, para capturar, reproducir los registros de vídeo de múltiples cámaras o sensores de profundidad. Estos registros capturados pueden ser utilizados para procesar el movimiento capturado. IPi Mocap Studio analiza el registro y traduce el movimiento físico del actor en data que se le puede aplicar a uno o varios actores virtuales. Debido a que el sistema de captura de movimiento está dividido en dos partes, la de grabar el video y el de procesar el video, esto permite que el primer paso se pueda realizar con computadoras de bajo rendimiento o laptops, porque el software es gratuito, se puede conectar en un número ilimitado de computadoras, dando la flexibilidad necesaria a utilizar varias para grabar una sola sesión a la vez, mientras se pueden observar la imagen de todas las cámaras y controlarlas.

La versión más reciente de iPi Recorder (ver. 3.1.3.38, salió el 19 de Noviembre 2015), puede grabar hasta con 16 cámaras Playstation Eye o 4 Sensores de Profundidad como el Kinect original o el Kinect 2, a la vez puede conectarse con varios controladores de movimiento como el Playstation Move. La cantidad de controles Playstation Move que se pueden conectar son cuatro por adaptador Bluetooth, y la verdadera limitación es cuantos adaptadores se puedan conectar a la computadora, ya que el software en si no tiene límites. También puede procesar

sesiones que hayan sido grabadas con versiones anteriores del programa para que sean procesadas por las versiones más actualizadas de iPi Mocap Studio.

Por su parte el software iPi Mocap Studio, es la parte del combo de Software que no es gratuita, el programa tiene varias versiones:

- Exprés: procesa la información de un sensor de profundidad.
- Básica: procesa la captura de 2 sensores de profundidad o 6 cámaras.
- Profesional: procesa 16 cámaras o 4 sensores de profundidad.

iPi Mocap Studio es compatible con la mayoría de los programas de animación (AutoDesk o Blender) y formatos populares en el mercado, aunque también es compatible con programas para videojuegos como el Valve Source Engine y Second Life, sin necesidad que se ajuste el esqueleto en programas especializados como Motion Builder para hacer la transferencia de movimiento a ese formato.

Para procesar el video grabado se recomienda tener una computadora de alto rendimiento, pero no es necesario tener procesadores o tarjetas de video profesionales, ya que fue diseñado y optimizado para ser utilizados con procesadores y tarjetas de video orientadas para computadoras de uso casero. Recomiendan tarjetas de video que estén dirigidas al mercado de videojuegos en computadora, ya que éstas tienden a ser las de más alto rendimiento. Esto se debe a que el software de iPi Mocap Studio, a diferencia de programas más comunes, no procesa sus datos principalmente a través del procesador de la computadora (CPU), sino que de manera similar a los videojuegos procesan la mayoría de los datos a través de la unidad procesamiento grafica (GPU). La razón de esto, en parte se debe a la capacidad de procesamiento de GPU de una computadora es mucho más fácil y económica de mejorar y mantener actualizada. Pero la principal razón es que como el nombre lo sugiere (GPU) están diseñadas para procesar data en forma de imagen, y como la tarea principal de iPi Mocap Studio es procesar la imágenes del video y reconocer el patrón de la figura humana para así rastrear su movimiento a lo largo del video. La necesidad de tener que reconocer la silueta humana en cada cuadro para poder así rastrear el movimiento, puede hacer que el tiempo que tarda la computadora sea en algunos casos entre 1 a 3 segundos por cuadro. El usuario tiene la posibilidad de escoger las porciones del video que desean procesar, y así pueden escoger la sección de video que más les interesa.

Los sistemas sin marcadores se pueden considerar el próximo paso en la evolución, desde el punto de vista tecnológico, ya que no se basa solamente en la triangulación con que funcionan los sistemas ópticos tradicionales, sino que busca ir más allá de la mera triangulación, porque busca entender y reconocer patrones para poder funcionar. Es importante señalar que el desarrollo de sistema sin marcadores va un paso más allá, pues busca desarrollar un tipo de visión artificial para que la computadora sea capaz de identificar cuál es el lado izquierdo o el derecho de un actor, así como también reconocer y distinguir los reflejos y sombras creados por el material con que está hecho la ropa que pudiera confundir al sistema. Tecnológicamente hablando, el avance de la captura de movimiento sin marcadores es mucho más complejo y requiere un conocimiento de cómo se ven las cosas y cómo se procesa lo que se ve, en comparación a cualquier otro sistema; es un área de la ciencia poco explorada en la actualidad.

Restore a three-dimensional scene using multiple cameras is easy. As we see the pixels from different angles, using triangulation, we can reconstruct a three-dimensional image. The main difficulty - in that part of the technology, which is responsible for pattern recognition. The program needs to understand, where a person is and where any part of the body.

(Ilyin, 2010)

Uno de los campos que ha estado explorando esta tecnología para entender, no solo cómo se ve y se entiende lo que se ve, son las compañías que están desarrollando sistemas para carros autónomos, debido a que necesitan que la inteligencia artificial que maneja este tipo de carro sea capaz de distinguir su entorno, por lo cual han desarrollado la capacidad de rastrear objetos y personas de manera rápida y eficiente. En consecuencia, no es difícil imaginar que en un futuro, esta tecnología en desarrollo se extrapole a sistemas como el de iPiSoft que funcionan de manera similar, lo cual permitirá mejorar las capacidades y los tiempos de procesamiento de sistemas de captura de movimientos sin marcadores.

Como se explicó previamente, iPi Mocap Studio procesa el movimiento de los actores, reconociendo los patrones que forman la figura humana, debido a esto no necesita ropa especializada, pero esto no significa que funcione con cualquier indumentaria, existen ciertos tipos de ropa que favorece la captura de movimiento. Para facilitar y evitar que haya errores se recomienda utilizar ropas que no sean tan holgadas, es decir, faldas o chaquetas anchas son poco



recomendadas, ya que confundirían y dificultarían la capacidad del programa de capturar el movimiento de manera precisa. Por eso recomiendan el uso de ropa más ajustada que facilite la definición de la silueta humana, un ejemplo de esto serían camisas manga larga, suéteres, o franelas, y pantalones de colores oscuros para evitar las sombras, especialmente en casos donde la iluminación no es uniforme.

De esta misma forma, hay algunos tipos de movimientos que son capturados con mayor facilidad que otros, un ejemplo de esto son los movimientos exagerados, como bailes o las artes marciales. Aunque pueden realizarse con facilidad movimientos normales como el caminar de una persona; sin embargo se tiene más dificultad con movimientos más sutiles, especialmente con los movimientos de las manos. Las versiones más recientes del sistema son capaces de capturar el movimiento de dedos. Todavía este sistema no puede procesar captura de movimientos faciales, ya que esto requiere equipos más especializados que los que están ellos actualmente trabajando, aunque se espera que en el futuro se tenga esta “característica”. Este sistema tiene también problemas al capturar movimientos cuando el actor está con su cuerpo entero en contacto con el piso.

#### 2.2.1.3.- Requerimientos para correr el software.

Ambos programas pueden correr en casi cualquier computadora que tenga un procesador Intel, que sea al menos Pentium 4 o más avanzado, o su equivalente fabricado por AMD. Aunque si es recomendable que se utilicen computadoras que sean Quad Core, especialmente cuando se quiera procesar datos de captura de movimientos prolongados. Requiere que la computadora esté corriendo un sistema operativo Windows Xp o más actualizado. Deben tener como mínimo 2 controladores USB, que estos sean USB 2.0 o USB 3.0; la computadora debe tener un sistema de almacenamiento, que tenga capacidades de escribir a más de 55Mbyte/sec y entre más grande mejor, ya que los archivos de data pueden ocupar varias gigas. Se debe tomar en cuenta que además de los requerimientos de hardware, hay unas recomendaciones de algunos accesorios, aunque no obligatorios que facilitarían el uso del software. Como extensiones USB 2.0, al igual que extensiones eléctricas, así como trípodes, al que se le pueda acomodar un Kinect. En cuestión de espacio el Kinect tiene la ventaja sobre otros sistemas, en que se puede utilizar en espacios relativamente pequeños, necesitando un área de 9 metros cuadrados para realizar la captura.

Se suele indicar, según refiere Nikonov que, le preguntaron si su sistema estaba listo para competir con la tecnología actualmente disponible en el mercado, las que utilizan la mayoría de los grandes estudios alrededor del mundo. A lo cual él respondió lo siguiente:

Eventualmente sí, pero en la actualidad nosotros recomendamos que esta tecnología sea utilizada por pequeños estudios, que necesitan más flexibilidad a un bajo costo, ya que los grandes estudios necesitan procesar el video de captura en tiempo real, algo que nuestra tecnología en este momento no permite. Pero a medida que el poder de las computadoras aumente y nuestro sistema sea optimizado, este defecto será una cosa del pasado

(Avlastimova, 2010)

### 2.2.2. –Brekel3D

Es una compañía holandesa, fundada por Jasper Brekelmans, quien es especialista en la animación 3d y ha estado activo en la industria desde 1999. En la actualidad trabaja para Motek Entertainment. Ésta es una compañía especializada en la creación de personajes virtuales, utilizando la tecnología de captura de movimiento en tiempo real. Su propósito es llevar la captura de movimiento a las masas, al igual que IPiSoft. Por lo cual ha desarrollado una serie de programas para la captura de movimiento sin marcadores, basados en el uso de sensores de profundidad, utilizando principalmente el Kinect de Microsoft. Su tecnología se diferencia de la desarrollada por IPiSoft, ya que se enfoca en la captura de movimiento en tiempo real y la captura de movimiento facial. Brekel3D, como lo indica en su página web, al referir que, se enfoca más a la creación de pipelines y herramientas para la creación de personajes animados 3d, que van desde la creación de flujos de trabajo y el rigging para proyectos interactivos, basados en la tecnología Kinect a soluciones para proyectos de animación más tradicionales (“About me,” 2010).

#### 2.2.2.1.- Funcionamiento del software

El software de Brekel3D funciona de manera diferente a como lo hace iPiSoft, porque funciona en tiempo real y es capaz de hacer captura de movimiento facial. Esto se debe a que ambos programas funcionan con algoritmos diferentes, mientras que los desarrolladores del software de IPiSoft utilizan la data en bruto que provee el Kinect, manejando solo el SDK del Kinect por sus drivers, y el proceso de captura y procesamiento de la data están separados, lo cual les permite realizar captura de movimiento con computadoras que serían incapaces de

realizar captura de movimiento en tiempo real, de esta forma su software es más accesible. Por su parte, el Software de Brekel3D utiliza los algoritmos creados por Microsoft para el Kinect, los cuales son capaces de la captura de movimiento en tiempo real, pero con un nivel de precisión inferior al sistema de iPiSoft. Su software es un tipo plug-in para Motion Builder de Autodesk, un programa especializado para la animación, con diferentes herramientas para mejorar la animación, y la creación de animación, siendo con técnicas de captura de movimiento o por Keyframe (“Pro Body – FAQ,” 2012). Todos los plug-ins de captura de movimiento que ofrece Brekel3D utilizan un solo sensor Kinect, sea el original o el Kinect de Xbox One, aunque están desarrollando un programa para usar múltiples sensores Kinect. Los diferentes plug-ins, aunque son compatibles entre sí, no se pueden correr de manera simultánea en una sola computadora, pero se puede grabar en diferentes computadoras, y luego integrar la data en una sola, aunque este proceso es manual, ya que hasta el momento no hay una forma automatizada de realizarlo (“Pro Body 2 – FAQ,” 2014).

#### 2.2.2.2.- Captura de movimiento facial

El software de captura de movimiento facial de Brekel3D, lo que lo diferencia de las otras opciones de software de bajo costo sin marcadores es el ofertar la posibilidad de rastrear los movimientos faciales, pero no tiene las mismas capacidades del software de nivel profesional, como indica su creador, una de éstas es que es incapaz de rastrear el movimiento de los ojos. Pero tiene la calidad suficiente para que se pueda aplicar en una gran diversidad de proyectos animados, además de tener la ventaja de no necesitar entrenamiento para entender los gestos del actor, como es necesario con los programas profesionales (“Brekel Pro Face,” 2012).

Dependiendo de la versión del software y cuál de los 2 sensores se utilice, los puntos de rastreos cambian, si se utiliza el Kinect del Xbox 360, se tienen 11 puntos de rastreo, mientras que si se utiliza el Kinect del Xbox One, se tienen 20 puntos de rastreo, los cuales incrementan las capacidades de animación; además que es 3 veces más preciso que el Kinect original (“Brekel Pro Face 2,” 2015). La data capturada con el software de manera similar a iPiSoft, se puede guardar en una gran cantidad de formatos. Se recomienda que, de manera similar a los otros programas de esta compañía, se tenga buena iluminación para evitar errores, o imprecisiones en la captura, y se puedan utilizar los filtros para refinar la captura.



## Capítulo III: La producción en la captura de Movimiento.

En el capítulo anterior, se precisó que para la captura de movimiento es necesario disponer, tanto de un equipo (hardware) con muchas capacidades como de un software adecuado, a fin de poder utilizar el equipo a su máxima capacidad. Así mismo, cabe recordar que, como en toda producción cinematográfica, es indispensable comenzar con la preproducción, la cual forma parte integral de la aplicación de la captura de movimiento; ya que esto refiere a cómo se integraría la captura de movimiento sin marcadores a los procedimientos de una fase de preproducción estándar. Por esto, puede decirse que de la misma forma que la preproducción es importante para la realización de cualquier proyecto cinematográfico, en el caso de la captura de movimiento, la preproducción es igualmente importante, ya que ésta permitirá arribar a un mejor resultado en la captura de movimiento.

Menache (2010) plantea que existe la creencia que la captura de movimiento es simple, muy rápida y que funciona de manera casi automática. Esto, desafortunadamente, no es cierto. Conseguir una buena captura, con datos que sean fáciles de utilizar, requiere paciencia y experiencia en el uso de sistemas de captura de movimiento (Menache, 2010). El proceso no será rápido hasta que el individuo o grupo de trabajo lo haya hecho varias veces y por ende, esté lo suficientemente experimentado en cómo funciona el sistema a utilizar. La captura de movimiento requiere de una planificación adecuada, tener un objetivo claro de lo que se quiere lograr, una idea clara sobre qué es lo que debe hacerse para poder lograr un proceso de captura de movimiento satisfactorio y un mejor producto. Por ello, es indispensable considerar los pasos que a continuación se describen.

### 3.1.- Preproducción.

La fase de captura de movimiento debe ser pensada desde un comienzo como parte integral del proyecto que se está realizando. Mucha gente se podría preguntar por qué se debe abordar desde la preproducción la captura de movimiento, bien sea si ésta sea realizada con los sistemas más tradicionales, como los sistemas ópticos con marcadores, o con los sistemas más modernos, los cuales utilizan sensores activos, como el caso del que se utilizó en este trabajo: un sistema sin marcadores. La principal razón de realizar la preproducción obedece a la necesidad de reducir el costo y tiempo de producción, así como también prever y evitar

problemas innecesarios, a fin de lograr como resultado un producto final de mejor calidad. El saltarse o realizar poca o nada de preproducción, es una garantía de enfrentarse a problemas en algún momento de la producción, lo cual causarán pérdida de tiempo, debido a que se tendrían que realizar más capturas de movimiento de las necesarias, podría también llevar a dificultades durante la captura o en el procesamiento de la data capturada, pues estos son algunos de los problemas que se pueden evitar o por lo menos reducir o atenuar durante la fase de preproducción. A continuación se presentan los pasos a considerar en la preproducción.

### **3.1.1. – Elaboración del Guión**

Como en todo proyecto audiovisual, sea éste una gran producción Hollywoodense o un corto estudiantil, surge a partir de una idea. Pero una idea por sí sola no hace real una película, es indispensable expandir esa idea y crear algo con ella; es decir se requiere elaborar un guion.

Un guion es la narración de la historia que se está tratando de contar. Esto podría ser cualquier cosa, aun cuando es preferible que se sigan los estándares de la industria. El guion es uno de los elementos básicos de cualquier tipo de proceso de animación o producción audiovisual. El guion permite llevar todas las ideas a una secuencia y le da un cierto flujo a la historia que se quiere contar. No sólo sirve como una hoja de guía para futuros planes, sino que también sirve como una base para el elenco. Este ayudara al elenco a prepararse intelectual y emocionalmente para su personaje, así como a entender lo que se requiere de ellos físicamente en la captura de movimiento (Field, 2005).

### **3.1.2. – Elaboración del Guion Grafico**

Después de completado el guión es recomendable hacer un storyboard o guión gráfico. Esto consiste en una representación visual del guión, por lo general es un conjunto de dibujos, con un breve texto en cada imagen. Cada dibujo representa un elemento esencial para la película, sea esto: en actuación, en la duración, la puesta en escena, la cámara (tamaño de la toma, movimiento de la cámara, ángulo de la cámara, etc.), edición (transiciones entre tomas, etc.), así como el sonido (de los diálogos, la narración, efectos de sonido, etc.) (Ver fig. 3.2).

Los dibujos para un guión gráfico suelen ser rápidos bocetos en miniatura, que constituyen representaciones básicas para pre-visualizar los movimientos y las emociones del guión. Los guiones gráficos han constituido tradicionalmente uno de los pasos más importantes

en la preproducción de cortometrajes animados, los estudios de Walt Disney los implementaron en la década de los 30, y fueron eventualmente adoptados por otros estudios para la producción de películas y programas de TV. Permiten a los creadores de contenido, directores, actores, cineastas o cualquier otra persona que está involucrada en la preproducción y en la producción a obtener una imagen y comprensión clara de cómo se verá el producto final. (Glebas, 2008)

En la actualidad hay muchas formas para crear un guión gráfico, hechos a mano para ser pegados en corcho, a programas especializados, lo importante es que se pueda agregar, eliminar, cambiar y reorganizar cualquier parte del guión gráfico, mientras se esté en la etapa de encontrar la mejor manera de contar la historia. Lo importante no es cómo se hace el guión gráfico, el objetivo principal y esencial del guión gráfico es aclarar, afinar y reafirmar las ideas, a objeto de poder comunicarlas de una manera visual al equipo de artistas y a los encargados de tomar decisiones (directores, productores). Un buen guión gráfico, es capaz de comunicar claramente las ideas principales de la historia.



Fig. 3.2 Muestra de varios cuadros de un guión gráfico en bruto, ya que carece de descripción de las tomas y los diálogos.

### 3.1.3. – Decidir si la captura de movimiento es adecuada.

Después de tener un guión completo, que haya sido revisado y acordado que es la versión definitiva, se pasa a la parte más importante en la preproducción, en el caso que se desee utilizar captura de movimiento. Eso es, decidir si la captura de movimiento es la tecnología correcta a ser utilizada. Ya que la captura de movimiento es una herramienta, que está a la disponibilidad del Artista/Creador, más que la solución para todas las situaciones en la cual esté involucrada la animación. Hay situaciones en las cuales el uso de la captura de movimiento no es la más adecuada ni recomendable, principalmente cuando se quiere un movimiento más estilizado y poco realista, la misma es más adecuada a formas más tradicionales de animación como la animación Keyframe o de pose a pose. La otra situación es cuando uno no tiene la experiencia suficiente operando sistemas de captura de movimiento, y el proyecto tiene una

fecha límite muy cercana, también cuando el presupuesto es muy ajustado y no permite corregir algún error.

It is safe to assume that in most cases performance animation will not save time, except when used to create a product with expected quality trade-offs ... Similarly, it is safer to believe that it won't save you money. I'm not saying that it is impossible to save time and money using motion capture, but that you should never assume it will unless you have demonstrated this by testing. Testing, of course, will cost time and money.

(Menache, 2010, p. 75)

Menache es un creyente de que la razón para utilizar la captura de movimiento debe estar basada principalmente en la estética del movimiento; es decir, que si se desea que el movimiento del personaje sea realista en base al movimiento humano, entonces uno debe utilizar la captura de movimiento. En este sentido, la visión de Menache es la más acertada, ya que como se había dicho antes, la captura de movimiento es una herramienta más a la disposición de los creadores, y debería ser escogida preferiblemente por sus cualidades estéticas. Aunque desde que salió el libro de Menache han salido alternativas que son más económicas para el uso de la captura de movimiento, como los sistemas de captura sin marcadores, pero esto no hace menos válida su recomendación; especialmente cuando tiene que ver con el tiempo, porque si hay casos cuando la captura de movimiento puede ser considerablemente más rápida que las técnicas de animación convencionales, pero esto requiere un nivel de experiencia amplio con la técnica de captura de movimiento.

I believe that a decision to use motion capture has to be initially based on look alone. If the realistic look of human motion is not what you want, do not even consider motion capture. This is the first bridge that you should cross.

(Menache, 2010, p. 76)

Según Menache, hay ciertos principios de la animación que no pueden ser logrados en su totalidad, o con cierta dificultad, utilizando la captura de movimiento, mientras hay otros que se prestan a la captura de movimiento. En la Tabla 3.1 que a continuación se inserta, se pueden apreciar los Principios inherentes a la captura y los difíciles o imposibles de capturar.



Tabla 3.1 Principios inherentes a la captura y los difíciles o imposibles de capturar.

Principios inherentes a la captura	Principios difíciles o imposibles
Arcos	Estirar y encoger
Acción superpuesta	Anticipación más allá de los límites físicos.
Movimientos Secundarios	Exageración más allá de los límites físicos.
Acción directa	Acción continuada más allá de los límites físicos.
Entradas lentas y salidas lentas	

Fuente: (Menache, 2010)

Si en la historia que se planteó en el guión y el guión gráfico se requiere movimientos que formen parte de alguno o varios principios de la animación que se consideran difícil o imposible con captura de movimiento, se debería considerar desde este momento que la captura de movimiento no es la herramienta más adecuada. Esto no significa que no se debería aplicar la captura de movimiento de un actor humano a un personaje, hay momentos en que se puede hacer, especialmente si uno busca el antropomorfismo, como en la película de Happy Feet, la cual se realizó con captura de movimiento, aunque los personajes son pingüinos caricaturescos, allí no se basaba en movimientos realistas de los pingüinos. El enfoque principal de la película está en las actuaciones musicales, el hecho de que los pingüinos vivan en una sociedad humanizada, donde cantan y luego terminan bailando ayuda a que el efecto de la captura sea convincente. Hay otras películas en las cuales se utiliza también captura de movimiento para animar personajes no humanos, que van desde los personajes humanoides en Monster House hasta la actuación de Benedict Cumberbatch, como el dragón Smaug en la película The Hobbit: The Desolation of Smaug. Esta escena del dragón se pudo lograr con la mezcla de la data que se obtuvo a través de la captura de movimiento, métodos como el pose a pose y la programación de procedimiento que permitieron aplicar y expandir lo conseguido con la captura de movimiento para que éste se vea como lo que nos imaginamos que es el movimiento de un dragón. Con ello se pueden utilizar los mismos métodos que permiten facilitar el uso de los principios de la animación como: estirar y encoger, exageración, y el uso delimitado de animación de pose a pose. Pero esto lleva a preguntarnos ¿Por qué hacer captura de movimientos realista, si quieres un estilo caricaturesco? En muchos casos la modificación manual de los datos de movimiento capturado, puede ser más costoso, tanto en tiempo o dinero que irse con un método, como el de pose a pose o animación de cuadros claves. Mientras que si se hace con los métodos de procedimientos programados no siempre son los más adecuados, ya que los resultados no siempre son precisos y debido a esto se deben utilizar con cautela.

Por esto es que siempre debe experimentarse antes, y ver si la idea que se tiene para lograr el estilo que se desea opera, porque siempre es más seguro ir a través de un proceso de prueba, a fin de saber si se va a producir el resultado deseado antes de comprometerse con un modo de trabajo, que puede luego llevar a un gasto alto o uso de mucho tiempo.

En este punto de la preproducción, cuando ya se ha decidido que si se va a utilizar la captura de movimiento. Asumiendo que la decisión haya sido utilizar captura de movimiento, el próximo paso a seguir sería determinar qué tipo de sistema de captura se utilizara, según las necesidades del proyecto. En muchos casos la decisión ya está tomada, ya que se posee un sistema de captura de movimiento, y en ese caso sería más bien ajustar el flujo de trabajo para que se aprovechen las ventajas del sistema que se dispone. Si es el caso que no se poseen los equipos, se pueden considerar cual alquilar o comprar, basándose en las necesidades del proyecto y cuáles son las ventajas de cada sistema, así como si se percibe que son las más adecuadas. Las ventajas de los diferentes sistemas fueron discutidas en el primer capítulo. Esta decisión varía entre proyectos, aquí se enfatiza en el sistema de captura de movimiento óptico sin marcadores, utilizando como hardware el Microsoft Kinect y como software iPiSoft iPi Mocap Studio.

Menache plantea que existen principalmente dos tipos de proyectos que se encuentran en las producciones audiovisuales, los proyectos lineales y los proyectos no lineales. Cada uno de estos determina como se procederá en el proyecto.

- Los lineales: son los proyectos donde las secuencias de acciones y movimientos han sido predeterminadas por los creadores, y no hay ningún tipo de interacción con el espectador. Ejemplo: los cortos, videos o películas.
- Los no-lineales: son los proyectos interactivos, donde el espectador tiene el rol de usuario/jugador, y controla los movimientos y acciones de uno o varios personajes, siendo éste el que determina la secuencia de eventos que ocurren en tiempo real.

Estos proyectos se abordan cada uno de manera diferente, pero pasa por las mismas fases de preproducción y desde este punto en adelante se trata como si se fuera un proyecto lineal, aunque las etapas de preparación de estos son similares.

If a character design is already available, and it has been determined that realistic motion is the way to go, it is time to evaluate if it is possible to capture the data you are looking for. For this purpose, I like to use a divide-and-conquer approach, breaking the project into small, similar groups of shots and using these categories to evaluate strategy and plan the sessions. A shot is an uninterrupted camera take that is later combined with other shots to form a scene. It will fit our purpose to define a shot as an uninterrupted piece of character motion, because this definition applies to all media, including linear and nonlinear projects. A step-by-step explanation of how to perform the evaluation follows.

(Menache, 2010, p. 83)

#### 3.1.4. – Lista de tomas

La lista de tomas es un documento que detalla y describe las tomas que conforman las escenas, las cuales en su parte conforman al film en su totalidad. El objetivo principal es que cada parte del guión tenga por lo menos una toma que la representa. Tiende a seguir el mismo orden que el guión, no el orden de grabación. Una lista de tomas tiende a incorporar los siguientes elementos: personajes en la toma, tipo de toma, movimiento de cámara, acciones importantes y objetos importantes. Esta es una de las etapas más tempranas y más importantes en la organización de un proyecto cinematográfico, junto con el guión gráfico; esto le permite al director y director de fotografía visualizar lo que se quiere en la toma, y desde ahí tomar las decisiones logísticas. La lista de tomas cumple múltiples funciones durante la producción, desde ayudar a planear la logística, servir como base referencial para el director y el director de fotografía, junto con sus asistentes, a objeto de asegurarse que durante la grabación no falte hacer ninguna toma. En la captura de movimiento, esta etapa no es diferente, solo que el enfoque va más allá de las tomas y se orienta en las acciones y movimientos de los actores, y como se vinculan con las tomas que se animaran con los datos capturados. Según Kitagawa y Windsor, la lista de tomas en la captura de movimiento es una lista de las acciones y movimientos que crean una escena. Ya que separar la historia en tomas, da una idea de lo complejo que puede ser: tal vez se necesite que el talento tenga entrenamiento especial, así como la utilería necesaria en cada toma (Kitagawa & Windsor, 2008).

En la cinematografía se connota como toma a lo que una cámara graba sin interrupción, mientras que la acción se lleva a cabo en el escenario, cuando la cámara deja de grabar, la toma termina. Una escena normalmente puede estar compuesta por múltiples tomas que representan

un evento continuo. Esto es algo que se tiene que tomar en cuenta cuando se planifican las sesiones de captura de movimiento, ya que se puede lograr la sensación de movimientos que serían humanamente imposibles, con el uso de edición y planificación de tomas, esto también facilitaría la etapa de post-producción. Un ejemplo de esto lo plantean Kitagawa y Windsor (2008). Si uno tuviera una toma en la cual se necesita que el personaje corra 45 segundos continuos, es casi imposible que se tenga un estudio de captura de movimiento para capturar ese tipo de toma, ya que el área es muy grande. Pero esto se puede recortar en movimientos más pequeños que se puedan interconectar entre ellos, o que se puedan convertir en un loop o bucle; es decir que el final del movimiento coincida con el inicio, y que el movimiento entonces parezca infinito.

Estos son los tipos de problema que uno tiene que tratar de resolver antes de ir al estudio de captura de movimiento, ya que facilita la aplicación de la data capturada, y ahorra tiempo y recursos al no tener que resolverlos el día de la grabación o en post-producción. Kitagawa y Windsor (2008) tienen una serie de preguntas que uno debería hacerse cuando va hacer captura de movimiento.

Are you going to be tracking facial, finger, and full body data all at once? If so, is this possible throughout the entire space or is it just possible in a smaller area of the space that has more cameras? What are the physical limits as far as length, width, and height of the motion capture volume?

(Kitagawa & Windsor, 2008, p. 16)

### **3.1.5. – Motion Capture Blueprints o Planos de Captura de Movimiento**

Menache (2010) recomienda que más allá de la lista de toma en la captura de movimiento, se debe hacer Motion Capture Blueprint o un plano de captura de movimiento. Estas es una mezcla entre el guión gráfico y una lista de tomas que se utilizarían en proyectos cinematográficos más convencionales, pero tienen información adicional que es pertinente a la captura de movimiento. Se necesita saber con qué esta interactuando el actor, desde su relación con el área de captura, otros personajes y utilería. También es necesario conocer la ubicación de cada una de estas interacciones en el espacio, y si el suelo es irregular, plano o si tiene alguna pendiente; de la misma forma si el actor se inclina contra la pared o si tiene que moverse sobre

un obstáculo. La importancia de los factores anteriores tiende a ser ignorada durante la fase de preproducción, ya que se suele pensar que esto no interferirá con el proceso de captura. Pero estos aspectos deben ser tomados en cuenta, ya que evita que haya complicaciones innecesarias en la etapa de captura. Otros aspectos importantes son las medidas de las diferentes piezas de utilería y si éstas pueden presentar problemas de oclusión o si el volumen del área de captura es lo suficientemente grande, lo cual puede facilitar la selección del estudio de captura de movimiento o escoger la locación para la captura.

Los planos para la captura de movimiento pueden ser creados en una base de datos, hoja de cálculo o software de gestión de proyectos. No importa cuál es el método para la creación de la misma, pero se deben incluir los siguientes elementos:

- Descripción de la toma:  
Una pequeña descripción de lo que pasa en la toma.
- Nombre referencial:  
El nombre por el cual se puede referir a la toma (si aplica)
- Nombre de los personajes:  
Una lista de los personajes que van a estar en una misma toma, y si su movimiento será capturado de manera simultánea o separado.
- Duración de la toma:  
La captura de movimiento puede ser costosa, especialmente si uno está alquilando un estudio de captura por tiempo. Por eso es recomendable saber cuánto tiempo va a durar cada toma, a fin de poder facilitar el cálculo del tiempo de captura. Menache (2010 ) también recomienda incluir handles, los cuales son unos momentos antes y después de la toma, especialmente si el movimiento lo requiere, un ejemplo de esto es cuando uno quiere lograr grabar un salto largo, en esos casos debería incluirse en el tiempo requerido para la toma también el tiempo que consume en correr para ganar la velocidad antes de saltar, ese tiempo previo sería el handle, y se puede decir lo mismo del tiempo necesario para parar el movimiento. Esto ayuda a calcular el tiempo total necesario para lograr la toma deseada y evitar irse o sobre pasar el presupuesto.

- El área necesaria para la captura:

Es muy importante conocer las limitaciones de volumen y el área de la superficie de captura de movimiento. La mayoría de los sistemas tienen un volumen máximo que no puede ser excedido, es necesario asegurarse de que cada toma cumpla con esas limitaciones. Hay algunos sistemas que permiten tener algunas libertades con el volumen, reduciéndolo en un sentido para aumentarlo en otro, pero sin exceder el límite total del sistema. Por ejemplo, la captura de un salto vertical requiere tener una altura mayor que un salto largo, mientras que el último requiere que éste sea más largo que ancho.

Incluir el área necesaria para las acciones es fundamental para poder catalogar las tomas, según el área necesaria y la configuración de las mismas. De esta forma la etapa de captura será más eficiente, ya que no se pierde tiempo recalibrando o reconfigurando el sistema, especialmente si se trata de un sistema de captura de movimiento con marcadores. Esta son unas ventajas de todos los sistemas ópticos y electromagnéticos, ya que se pueden cambiar el posicionamiento de las cámaras o sensores para mejorar el área de captura. Si las diferentes configuraciones son planeadas con antelación los cambios se pueden hacer con la menor frecuencia posible, y pueden ser realizados de manera rápidas, ya que los posicionamientos de los equipos fueron pre planeados previamente (Menache, 2010).

Aunque no siempre es necesario tener que reconfigurar el set de captura de movimiento, hay casos en los que se requiere tener más de un set de captura de movimiento, cada uno con diferentes configuraciones listas para funcionar. Aunque si es válido decir que, estos casos tienden a ser excepcionales, y reservados a grandes producciones. Menache (2010) menciona que la película *The Polar Express* (Zemeckis, 2004), se realizó en tres distintos sets de captura, los cuales fueron alquilados de manera simultánea, a fin de facilitar la transición rápida entre los sets. El primer set era el más pequeño de todos, cuyo volumen era un cubo con 3m por lado. Era utilizado para todos los primeros planos y planos de capturas faciales. Alrededor del volumen de captura, se tenían posicionadas 70 cámaras con lentes tele, ya que se necesitaba el alcance para cubrir todos los marcadores faciales y otros 10 con lentes angulares, porque se necesitaba el plano más amplio para cubrir los movimientos del cuerpo entero. El segundo set tenía un volumen mayor, era de 9 metros de largo por 18 de ancho y 3 de alto, se utilizaba para capturar tomas amplias y actuaciones

que necesitaban desplazamientos por áreas grandes, como un largo paseo o muchos personajes que interactúan en un espacio amplio. El tercer set tenía un volumen de captura más alto, un cubo con 9 metro en cada lado, de manera casi exclusiva para las acrobacias que requerían tramoyas especiales (Menache, 2010). Los volúmenes segundo y tercero tenían alrededor de 100 cámaras cada uno. Aunque la grabación en los diferentes set nunca se realizó de manera simultánea, este tipo de producción salía más económica, porque no tenían que mantener parados al reparto ni al equipo de producción, como sería en el caso de tener que reconfigurar para las diferentes tomas.

- Configuración del Personaje:

La configuración del personaje, no se debe confundir con la configuración de marcadores en los personajes. En este caso, se refiere específicamente a la configuración del personaje virtual, y no tanto en su apariencia física o con el vestuario, sino más en su rigging, que en el mundo de la animación se refiere al proceso de configurar los huesos y la localización de las articulaciones, este proceso determina los movimientos y las posibles deformaciones del cuerpo cuando se mueve. En la actualidad hay una variedad de software disponible en el mercado, desde las opciones pagas como *Maya* y *3D Max* hasta las opciones libres como *Blender*. La configuración del personaje depende de la configuración de los marcadores, ya que los datos recogidos deben ser suficientes, a objeto de calcular toda la información necesaria para animar las articulaciones. Por ejemplo, si se tiene una articulación de la rodilla en la configuración del personaje, se le pueda añadir los marcadores necesarios para cuando se haga la captura de movimiento y se pueda calcular las rotaciones de la rodilla.

No es necesario determinar el diseño definitivo de la configuración de los marcadores en esta etapa, si el sistema que necesitas los requiere, pero se debe tomar en cuenta, ya que si la configuración del personaje no es compatible con la configuración de marcadores posibles, se debe rehacer la configuración del personaje. En el caso del sistema sin marcadores, como el tratado en el capítulo anterior, donde el software de *iPiSoft*, utiliza un esqueleto o rigging pre-diseñado para poder calcular la captura, ya que necesita rastrear la silueta humana sobre marcadores, porque uno debe configurar al personaje con el que sea compatible. Si el sistema sólo puede manejar una configuración de marcadores, como es el

caso del sistema sin marcadores, o los sistemas con trajes, la configuración del personaje tendrá que ser diseñado sobre la base de esta restricción de configuración.

- Eje Global

Es importante mantener un eje global a lo largo del proyecto, especialmente si se necesita capturar la data de varios personajes que interactúan entre ellos. Esto asegura que las interacciones siempre concuerden, y evita perder tiempo ajustándolas manualmente en post-producción. Debido a esto Tobón (2010) asegura que siempre cuando se vaya a comenzar una sesión de captura de movimiento, uno se debe asegurar que la orientación y la posición del eje global en el programa de captura de movimiento concuerden con el eje del programa de animación y sean compatibles (Tobon, 2010).

- Unidades de medida

Después de establecer el eje global para la producción, es bueno establecer las unidades de medidas con la cuales se trabajará. Esto facilitará crear una escala en caso que los personajes sean más grandes que los actores, y también permite la aplicación de la data capturada en el mundo virtual creado para la producción. Por esto es muy importante que las medidas estén basadas en medidas del mundo real, sean estos metros, centímetros, pulgadas o pies, y sobre medidas que solo existan en la computadora.

- Las cámaras, los ángulos y el encuadre.

Conocer la posición de la cámara, el ángulo de la misma es importante en todas las producciones audiovisuales, pero tiende a ser olvidado en la captura de movimiento, ya que no se está grabando la actuación final. Sin embargo, es importante tomarla en cuenta por dos razones, especialmente en proyectos lineales. En primer lugar, cuando se captura la actuación siempre se quiere tener una referencia grabada. Si el proyecto es lineal, siempre es mejor tener al menos una cámara colocada en el ángulo y el encuadre final de la cámara, a fin de ser utilizado como referencia al lado de la data capturada al momento de post-producción. La referencia con la grabada será utilizado en algunos casos para decidir entre las tomas, o como un marcador de posición para fines de edición, y, definitivamente, como referencia a la hora de aplicar los datos al personaje final. En segundo lugar, el intérprete debe saber dónde está la cámara, cuando se actúa, se hace para la cámara, que es donde está



el público. Si no se sabe dónde está el público, no se puede establecer contacto con él a través del lenguaje corporal y la actuación se vuelve plana. Si no es posible colocar la cámara de vídeo en la dirección general de la cámara final, es necesario dejar que el actor sepa cuál es la posición final de la cámara.

La posición de la cámara no debe ser confundida con el eje global de la moción. La ubicación de la cámara puede cambiar entre toma y toma, pero todas las tomas deben mantener un eje global común.

- Detalles de las interacciones:

Cada vez que un personaje interactúa con otros personajes o con parte de la utilería, debe ser detallado, ya que así se puede determinar cuál será la mejor forma de configurar el área de captura para evitar la oclusión de algún marcador durante la captura, y de esta forma evitar el uso excesivo de la revisión manual de la data capturada durante la aplicación, ya que son normalmente más necesarias durante las interacciones que en cualquier otro tipo de movimiento.

- Utilería, con tamaño y locación:

Tras saber cuáles serán las interacciones que tendrán el actor o actores a lo largo de la toma, se debe saber cuáles son las piezas de utilería que estarán en el área de captura durante la toma, considerando tanto su tamaño como su locación. Esto se debe basar en las unidades de medidas establecidas en el proyecto y siguiendo el eje global que se determinó para el proyecto. Todo esto, asumiendo que el proyecto esté bien organizado, y que se haya cumplido el paso de establecer la unidad de medición que se hayan previsto para el proyecto, además que las medidas del personaje sean proporcional a las del actor; es entonces cuando el posicionamiento de la utilería en el área de captura se pueden traducir y emular en el escenario digital con precisión, y de esta forma ser muy similar o igual al set real. Dependiendo de la complejidad de la interacción, se debe decidir el nivel de detalle de la utilería. Si por ejemplo, la utilería representa una columna que el actor debe evitar mientras corre, ésta puede ser demarcada con una marca en el piso, si el actor debe tocarla, ésta puede ser un tubo delgado o malla que marque la superficie externa de la columna virtual. Pero

uno debe tener presente siempre el evitar que las piezas de utilería causen oclusión de los marcadores, ya que pueden causar más problemas de los que resuelven.

En los sistemas que no son capaces de la captura en tiempo real, como es el caso de la mayoría de los sistemas ópticos, es imprescindible el uso de la utilería, ya que no se puede tener una referencia digital, como lo es en los casos con los de captura en tiempo real. Por consiguiente, cuando la captura es en tiempo real, se puede prescindir de la utilería, ya que se puede guiar y mejorar su actuación, porque se pudiera ver en tiempo real a través de diversas pantallas la actuación en el escenario digital, y ello con tan solo ver una de las pantallas. Además, esto facilita la producción, ya que no se necesita la construcción de utilería física para poder realizar las escenas, excepto en casos que haya una interacción directa.

- Tramoya

En algunas ocasiones los actores, sin importar que tan hábil sean, no pueden hacerlo por su cuenta, en estas ocasiones se necesita una especie de tramoya para realizar el movimiento deseado. No se les debe considerar utilería, ya que éstas no formaran parte de la escena digital. Un ejemplo de esto sería un sistema de cables con arnés para imitar que el actor está saltando más allá de las posibilidades humanas, o para simular como si estuviera volando. Esto se debe tomar en cuenta, ya que en algunos casos puede que la tramoya oculte parte del actor y la data capturada no sea la más precisa. Por consiguiente, es indispensable tomarlo en cuenta en el caso que se tengan que realizar cambios al sistema.

- Formato para los nombres de archivos.

Cuando se trata con archivos digitales, sean estos de sesiones de captura de movimiento o de grabaciones hechas con cámaras digitales en producciones más convencionales, uno tiende a lidiar con muchos archivos digitales. Los cuales pueden perderse en la desorganización con mucha facilidad, por eso es recomendable tener una serie de reglas, a objeto de poder de esta forma nombrar o identificar los archivos de forma adecuada y facilitar tanto su uso como su organización.

- Diálogo

El diálogo, tiende a incluirse cuando se habla de captura facial, pero Menache (2011) recalca la importancia de tener también el diálogo completo en las sesiones en donde se hace captura de cuerpo entero, porque cuando uno habla en la vida real, el lenguaje corporal es importante en la comunicación, porque puede perderse la inmersión y el realismo cuando no se tiene presente. Por lo tanto, si una toma tiene diálogo, es importante tenerlo en cuenta durante la captura de movimiento de cuerpo entero, esto debido a que puede alargar el tiempo de la actuación y la actuación en sí.

Los matices que son posibles de capturar, pueden ser tan realistas que si se omite un elemento tan simple como éste, el desempeño final no se verá creíble. El ojo humano está tan bien entrenado para las cosas reales que, si hay una pequeña desviación de lo natural anulará el efecto entero.

- Tomas a ser combinadas:

Cuando la toma es parte de un ciclo o un bucle, esto se debe especificar, ya que el inicio de la toma debe mezclarse con el final de sí mismo y si es un bucle con el final de la toma que la precede en el ciclo. Las posiciones de inicio y fin deben ser especificados en este tipo de tomas, preferiblemente mediante dibujos o fotos. Los proyectos lineales, por lo general tienden a no tener que mezclar las tomas para lograr los movimientos, ya que este aspecto es utilizado principalmente para proyectos no lineales, pero hay casos en que si es necesario mezclar las tomas para tener movimientos más grandes y dramáticos, especialmente si el área de captura que se dispone no es muy grande. Esto no es necesario digámoslo así, si el guión técnico enumera dos o más tomas que representan el mismo movimiento desde diferentes ángulos de cámara. Ya que éste es capturado en una sola toma y luego se posicionan las cámaras digitales para dar las diferentes tomas.

Para los proyectos no lineales, como los videos juegos, es probable que las tomas tengan que mezclarse con más de una toma y en algunos con muchas otras tomas. Esto se evidencia cuando un jugador controla lo que hacen los personajes, y el motor del juego reproduce el movimiento que corresponde a la acción del jugador en tiempo casi real, mezclándolo con

el movimiento precedente. Si estos dos movimientos no se fusionan así, el juego pierde fluidez.

- Cantidad de archivos:

A partir de este punto, el resto de las secciones del plano de captura de movimiento es para ser llenadas durante la etapa de captura, a fin de ayudar a mantener organizada la producción. Es en esencia la enumeración del total de archivos grabados que componen la toma. La situación más probable en la cual se pueda tener más de un archivo, es si hay varios personajes que necesiten ser capturados por separado, y si se necesitan varias tomas para que éstas sean combinadas o si un movimiento en particular necesita ser recogido en dos o más pasadas.

- La tasa o frecuencia de muestreo

La frecuencia de muestreo, según Menache (2010) es similar a los cuadros por segundo de las cámaras más tradicionales. Hay dos tipos de tasa de muestreo que se tienen que tomar en cuenta cuando se hace captura de movimiento, la primera es la frecuencia de captura y la segunda es la frecuencia de entrega.

- La frecuencia de muestreo de captura: es la frecuencia con la cual el sistema recolecta la data, en muchos casos, entre más alta sea la frecuencia ésta resulta en una mejor captura, aunque se debe balancear, ya que también aumenta el tiempo de procesamiento. Normalmente, uno no sabe con anticipación cual será la frecuencia que se utilizara con el equipo, por esto se debe dejar esta área en blanco para que se llene en el momento de la captura. Aunque se recomienda tomar en cuenta la frecuencia de muestreo de entrega para así facilitar la aplicación de la data capturada.
- La frecuencia de muestreo de entrega: es la que se debería tener en esta etapa de la pre-producción, ya que ésta es la que se aplicará a los personajes y se debe planear según el formato de distribución del corto o película que se prevé utilizar. Si el proyecto se va a representar en un formato de vídeo NTSC, la frecuencia de muestreo de entrega será igual a 30 (29,97) cuadros por segundo, mientras que para una película, lo más probable es que sean 24 cuadros por segundo. La frecuencia de entrega puede ser superior a la del formato final, pero siempre se debe tomar en

cuenta que ésta debe ser divisible entre ella, es decir si el proyecto es una película que se grabó en 24 cuadros por segundo, la frecuencia pudiera ser 24, 48 o 72 cuadros por segundo, lo cual puede dar la libertad para la creación de escenas en cámara lenta.

A pesar de esto se pueden presentar complicaciones en la aplicación de la data capturada, especialmente si se tiene que cambiar los cuadros por segundo, por esto siempre se recomienda aplicar la data a una animación incluyendo un timecode de sincronización. Para evitar cualquier posible desfases por las diferentes frecuencias de muestreo.

- Nombre de la configuración de marcadores:

Este es un campo que se deja abierto para ser utilizado, como se dijo al principio, este plano para la captura de movimiento, es una guía a utilizarse a lo largo de la producción, y debe permitir las anotaciones de cómo es cada uno de los casos. Ya que las configuraciones de los marcadores se hacen en la etapa de pre-captura. Esto es esencial para proyectos en los que se utilizan marcadores, especialmente si son proyectos grandes que requieren múltiples configuraciones. Es importante mantener un orden a lo largo de la fase de producción, a fin de saber que configuración de marcadores es utilizada en cada toma y así facilitar la postproducción, especialmente cuando se debe aplicar la data capturada con la configuración del personaje. En la captura de movimiento sin marcadores, se tiende a vincular la configuración de la silueta y su esqueleto, con la configuración del personaje y se guarda en un archivo, el cual se puede identificar con el nombre de la configuración de marcadores.

- Nombre de la calibración:

Cuando se recolectan los datos en las sesiones, estos deben estar organizadas para evitar confusión durante la etapa de post-producción; por consiguiente, se debe tener un campo abierto para poder vincular el archivo de calibración con las sesiones grabadas. Ya que si no se utilizan los mismos archivos como sería el caso de un sistema óptico, puede crear errores en la data y hacer que ésta sea imprecisa. Se debe anotar el nombre de la calibración a que se refiere cada toma. En el momento de análisis de datos, el operador tendrá que asociar la calibración con el archivo de datos brutos de captura para

que el software pueda comenzar el procesamiento de los datos. Este campo se deja en blanco en esta etapa, ya que sólo se utilizará en el momento de la recogida de datos

- Actores en escena:

Es importante utilizar siempre el mismo actor cuando se captura el movimiento para un personaje dentro de un mismo proyecto. Sería similar a cambiar actores en el medio de una película. Este campo se deja abierto hasta que se haya hecho la sesión de casting, donde se determinará el nombre o nombres de los actores que harán la captura de movimiento. Según la escena se pondrá el nombre del actor correspondiente en este campo.

- Instrucciones especiales:

Durante la producción pueden ocurrir ciertas peculiaridades, que deben ser recordadas para evitar confusión durante la post-producción, esta área está disponible para dar cualquier tipo de instrucción que no encaje en ninguna de las secciones anteriores. Su propósito es asegurar que toda la información es pertinente al proyecto y que la misma sea anotada y pasada de forma organizada a medida que el proyecto va avanzando.

### **3.1.6. – Realización del Animatic**

Según Kitagawa y Windsor (2008) el animatic es básicamente un guión gráfico en un formato basado en el tiempo que tiene elementos visuales, pero principalmente: bocetos animados, dibujos con movimientos de cámara y efectos especiales preliminares. También se incluyen elementos de sonido como: el diálogo, los efectos de sonido y la música. Ambos elementos transcurren en el tiempo en el que ocurrirían en la película, editándose de tal forma que ambos concuerden. Los animatic dejan ver el tiempo que duran las escenas y se ven los movimientos de cámara, para así poder determinar con facilidad si la historia se está contando con eficacia o no.

Conocer la duración de las escenas es esencial, no sólo en contar una buena historia, sino también en la conservación de recursos, sean estos monetarios o de tiempo, tanto de producción como de post-producción. Por esto, si se puede reducir momentos innecesarios en alguna toma, pueden ahorrarse estos recursos de la misma forma que se puede revisar la manera en que se

cuenta la historia, y poder hacer correcciones antes de que estos casos terminen siendo costosos. Un error común es tratar de capturar movimientos largos o complejos, sólo porque es posible y no porque sea favorable para la historia. Si no contribuye a la narración de la historia no se debería incluir, ya que puede terminar sobre cargada. No capturar cualquier escena que no contribuyen a la narración de historias. Un animatic ayudará a ver si el ritmo de la historia está funcionando, así como si todas sus escenas son realmente aporte a su historia en general.

Cada vez más los motores de videojuegos se están utilizando para crear animaciones. Esta es una herramienta muy útil, ya que la capacidad de representación en tiempo real de los juegos de video permite la actualización instantánea y sin cambiar los tiempos que tendrías con un animatic tradicional, hecho a mano o con el software de animación de largo renderización. Una buena biblioteca de movimientos genéricos para este propósito nunca es una mala idea.

En la actualidad muchos de las medianas y grandes productoras están utilizando tecnologías de videojuegos, especialmente los motores gráficos, que dan imágenes de alta calidad en tiempo real. También si se poseen, se pueden utilizar datos de otros proyectos donde se utilizó captura de movimiento, para así usarlos de referencia en el animatic. Es recomendado mantener una base de datos de movimientos genéricos, de modo que facilite la creación de animatic, si se realizan más proyectos de animación.

### **3.1.7. - El talento y audiciones**

La captura de movimiento, más que cualquier otro tipo de actuación, requiere que el actor sea capaz de realizar los movimientos con naturalidad. En el proceso de audiciones, siempre se debe buscar una persona que esté adiestrada en el área y que la practique de forma continua, de nada sirve buscarse a un actor que en su niñez aprendió karate o algún otro arte marcial, pero no lo ha practicado en los últimos años, cuando el personaje se supone que es un experto en artes marciales que las ha practicado toda su vida. Esto no es decir que si el actor tenga que poder hacer todo lo que el personaje requiere, si son escenas que ocurren con poca frecuencia, se pueden realizar las coreografías para que sean ensayadas hasta que sean naturales. No puede haber nada más dañino para un proyecto de captura de movimiento que tener una mala selección de talento, porque no importa que tanto se trate de corregir una captura que no sea adecuada con las expectativas que tiene, tanto uno como productor audiovisual o como

espectador, si el movimiento no parece natural para el personaje, no le gustara y puede hacer que el proyecto como un todo fracase. Si se está haciendo captura para personajes que son de fondo, cuyo rol en una película no animada sería realizada por extras, se debe evitar utilizar actores que tengan movimientos peculiares, ya que se terminaría teniendo varios personajes de fondo con movimientos similares, lo cual puede causar una distracción para la audiencia.

Por último, siempre se debe tratar de buscar talento que tenga la edad aproximada del personaje, si el personaje es una persona mayor o un niño, se debe tratar de conseguir actores que puedan sean personas mayores o niños, ya que es casi imposible poder conseguir que un actor o actriz en sus años veintes o treinta, se mueva con la misma naturalidad que personas que estén en sus 60s o 70s. Esto se debe, como se indicó en momentos anteriores, que hay movimientos inherentes que no percibimos cuando están presentes, pero si faltan nos parece poco natural y extraño.

Menache (2010) recomienda que, para las audiciones, si es posible utilizar de una vez, o preparar una prueba, donde tendrán que realizar ciertos movimientos que representarían los movimientos más importantes del proyecto. Si se puede realizar las audiciones con capturas de prueba facilita el proceso, debido a que los datos por sí solos revelan más sutilezas que no son evidentes cuando se mira en la actuación en vivo, al igual que las pequeñas particulares del movimiento de una persona específica que no serían aceptables para el personaje digital. Por esto, se recomienda que se escoja al actor, tras mirar solamente los datos, sin saber quién es el artista, ya que es probable que así se haga una mejor elección. Se pudiera pensar que es una gran pérdida de tiempo el tener que poner marcadores en cada uno de los solicitantes, pero el hecho es que la actuación se verá afectada, si el actor no se siente cómodo con los trajes de captura, por lo cual un ensayo general es necesario (Menache, 2010).

### **3.1.8. – Itinerario de captura:**

La fase final del proceso de preproducción es la creación del itinerario de captura, similar a los itinerarios de grabación en una producción tradicional. Basándose en todo lo hecho en las partes anteriores, especialmente en los planos de captura de movimiento. Se tienen que tomar en cuenta las categorías que son constantes y las que puedan ser variables. Las categorías constantes son aquellas que no se pueden cambiar, porque los actores o los equipos de tramoya



solo van a estar disponibles en esos días. Mientras que las variables, son aquellas que pueden realizarse en cualquier momento. Las categorías a que se debe dar preferencia, al momento de la programación y planificación del itinerario u horario a seguir son: el volumen de captura, los personajes, las interacciones, la utilería y tramoya, ya que son las que principalmente determinan si se puede o no realizar la captura. Lo que también se debe recordar es que, en la cuestión del volumen de captura si se necesita reconfigurar el sistema para lograr las diferentes tomas, se deben tratar de agrupar todas las tomas que utilicen dicha configuración. Debido a que reconfigurar un sistema toma tiempo, mientras éste se prueba y recalibra, limitando las horas en que se puede utilizar el set durante el día de configuración.

### **3.2. – Trabajo de pre-captura de movimiento**

Una vez pasada la fase de preproducción, donde se tiene el enfoque general y todos los aspectos de la producción, se pasa al trabajo de los aspectos que están más relacionados a la captura de movimiento. También hay un trabajo de pre-captura, que se realizan ya en las etapas finales de la preproducción y a lo largo de la etapa de producción, los cuales se deben seguir para tener la mejor captura posible, hacer que el proceso sea más ameno para los participantes y que se realice con la mayor efectividad posible.

Si se utiliza un sistema con marcadores, ésta es la etapa donde se planifican y se nombran los sets de marcadores que serán aplicados a cada miembro del reparto, si se desea saber más sobre este proceso, se recomienda revisar los libros de Kitagawa y Windsor (2008) y Menache (2010).

#### **3.2.1. – Preparar el área para la captura de movimiento:**

Cuando se va a realizar la captura, es muy importante la delimitación del área de captura para facilitar el trabajo, esto en la captura comprende:

#### **3.2.2. – Ensayos**

Los ensayos, según Menache (2010) son de vital importancia para el buen funcionamiento de una sesión de captura de movimiento. Por ensayos, nos referimos a que el actor(es) ensayará en un lugar fuera del área de captura. Existe muchas razones para tener ensayos fuera de las instalaciones, aquí se consideraran algunas de ellas. La principal razón, es la económica, pues en un estudio de captura de movimiento (mocap) el tiempo es dinero. Cada

vez que se inicia el uso de un estudio mocap, alguien tiene que pagar por la electricidad, los salarios del personal que trabaja allí, y el desgaste de los computadoras, cámaras, trajes mocap y marcadores para nombrar sólo algunos de los posibles costos. Resulta más económico ensayar los movimientos y así requerir menos tiempo en un estudio de captura, pero también permite ahorrar recursos y tiempo.

Ensayar ayuda a entender mejor lo que realmente se está buscando, permite conocer si se quiere algo nuevo que no se había pensado antes hasta que se ve delante de uno. Ensayar ayuda a crear una mejor lista de tomas. Es importante informar al personal con que se trabaja sobre el ensayo, esto ayuda a obtener ideas sobre cómo irá el rodaje mocap, la cantidad que posiblemente se puede capturar en un día, etc. Planificar por adelantado y pensar en todas las cosas que hagan el proceso mocap más fácil. Un ensayo puede ahorrar tiempo y dinero. En especial ayudará a obtener mejores resultados, porque todo el equipo entiende lo que se está haciendo y por qué hará también que su metodología sea mejor y el proceso será más agradable.

### **3.2.3. – Designar y preparar la utilería.**

Para esta etapa, la mayoría de la utilería que se va a utilizar a lo largo de la captura de movimiento debería estar lista, a fin de evitar cualquier tipo de retraso, se debe recordar que la utilería sirve como la contraparte analógica de una parte del set digital con el cual el personaje interactúa de manera directa o indirecta. Dependiendo del tipo de interacción puede ser altamente detallada la pieza de utilería o solamente una marca referencial en el piso. Siempre se recomienda que cuando se hacen los ensayos, se realice con la utilería, en especial, cuando se está utilizando sistemas ópticos de captura de movimientos, porque permite observar una posible oclusión de un movimiento, lo cual ayuda a tomar decisiones, bien sea en cambiar el movimiento o arreglarlo en la postproducción. Con algunos sistemas que tienen marcadores como los electromagnéticos, es importante tener en cuenta que la utilería de metal interfiere con la señal, esto indica que es necesario cambiar esa pieza de metal por otra. Ensayar con la utilería evita contratiempos y asegura que todo va a coincidir después.

## **3.3. – Captura de Movimiento**

En las sesiones anteriores nos enfocamos en cómo es el proceso para la preparación la captura de movimiento en general. De aquí en adelante, se especifica cómo es el proceso de

captura de movimiento con el sistema óptico sin marcadores, utilizando el Software de IPiSoft y dos (2) sensores de profundidad, en este caso el Kinect.

### 3.3.1. – Calibrar el equipo

La primera fase es la calibración del equipo, ésta es la más importante para poder comenzar la captura y es necesaria realizarla varias veces a lo largo del día, a fin de evitar errores y obtener la mejor data posible, Menache (2010) recomienda que uno debe calibrar al comienzo de la sesión y después de cada descanso prolongado (almuerzos), ya que algunas de las cámaras se pueden haber movido, aunque sea ligeramente durante esos momentos; también se debe recordar que, si en algún momento se movió algunas de las cámaras, debido a algún accidente, se debe recalibrar el sistema.

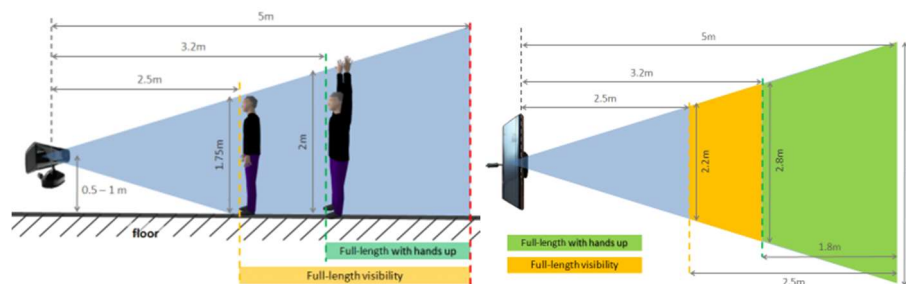


Fig.3.4 Volumen de captura posible por un sensor de profundidad Kinect Xbox 360

Cada sistema de captura se calibra de manera diferente. En este trabajo se enfatizó en la calibración del sistema óptico sin marcadores de iPiSoft, utilizando dos sensores de profundidad Kinect, el sistema puede utilizar hasta 4 Kinect de manera simultánea, pero la calibración es igual para un sistema que tenga 2 o más sensores. Al calibrar es necesario saber cómo posicionar los Kinect para poder sacarle mejor provecho al espacio, como se ve en la figura 3.4, el Kinect tiene un área muerta en los 2 metros y medio frente al sensor, ya que se le imposibilita capturar movimiento de un cuerpo completo, porque no puede ver el cuerpo entero, y tiene un rango máximo de 5 metros desde el sensor, lo cual nos da un área de captura un poco mayor a los 4.2 metros cuadrados, esta puede ser mayor si el actor es más bajo, o se puede posicionar el Kinect en posiciones superiores a un metro de altura, pero se sacrifica parte del volumen superior de captura; es decir que no se pueden mover los brazos muy altos sobre la cabeza. Otro aspecto a tomar en cuenta es que el Kinect puede tener dificultad rastreando, si se alejan mucho los Kinect ya que se puede exceder de manera accidental los límites de profundidad del sistema, se

recomienda marcar con tirro los límites más cercanos de captura (2.5m) y los más alejados (5m) para que el actor tenga una referencia física de los límites del área de captura. En el caso que nos ocupa existen 2 posibles configuraciones.

Primera configuración:

En esta configuración, los 2 sensores Kinect son colocados en un ángulo que sea entre 60 y 90 grados, y que cada sensor este entre 3 o 4.2 metros de distancia, como se puede ver en las figuras 3.5; esta configuración se recomienda cuando el personaje no tiene que hacer giros mayores a 180 grados, ya que ésta provee una mejor captura frontal.

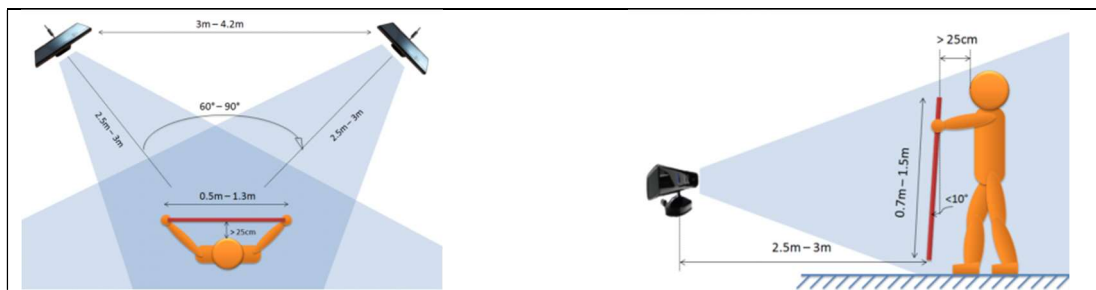


Fig. 3.5 Primera configuración con doble Kinect y tabla de calibración.

Para realizar la calibración se requiere una tabla rectangular plana, puede ser de cartón, madera o anime, lo más importante es que tenga las dimensiones correctas, debe tener un grosor máximo de 3cm, si es más grueso no se logra la calibración. En cuanto a su ancho debe tener entre 0,5 metros como mínimo, aunque el creador del software recomienda que sea de 1 a 1,3 metros de ancho y de largo debe tener como mínimo 0,7 metros, pero lo más recomendable es que sea entre 1 y 1,5metros. La toma de calibración se graba de la misma forma que una toma de acción; es decir, que el software de iPiSoft Recorder no tiene una función de calibración, sino que uno debe grabar la toma y luego designarla como la toma de calibración en el programa de Mocap Studio. En la figura 3.5, se puede observar la manera de sostener el cartón para realizar la calibración, y moverlo de un lado, girándolo lentamente de izquierda a derecha, evitando que el cuerpo tape uno de los sensores de profundidad; sin que se exponga la parte lateral a uno de los dos Kinects, como se puede ver en la figura 3.6, en la cual si se gira de manera exagerada se puede ver el cuerpo y el lado del cartón, lo cual hace que la calibración no se pueda concretar, ya que se pierden algunos puntos de rastreo.

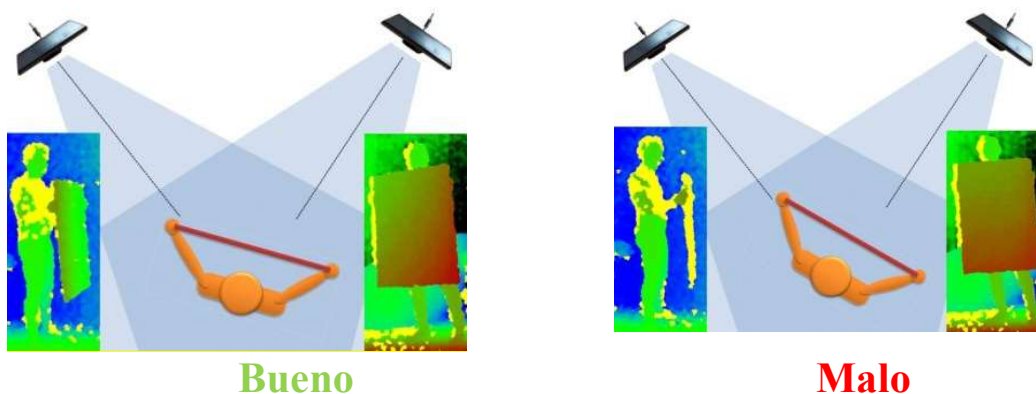


Fig. 3.6 Calibración del Kinect utilizando el software de iPiSoft con la primera de las dos posibles configuraciones con 2 Kinect.  
Segunda configuración

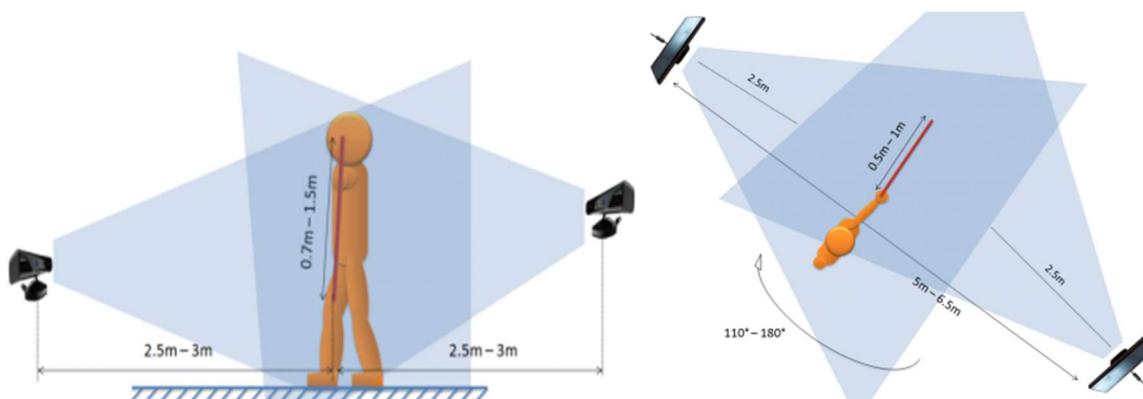


Fig. 3.7 Segunda configuración con doble Kinect y tabla de calibración.

Esta ocurre cuando los sensores están en un ángulo de  $180^\circ$ ; es decir, opuesto el uno al otro, lo cual facilita la captura o cuando se tiene que hacer giros de  $180^\circ$  y dar libertad para realizar movimientos acrobáticos (ver fig. 3.7). De manera similar que en la anterior configuración, en este caso se utiliza un cartón con las mismas dimensiones dadas anteriormente, pero en lugar de agarrarlo con los brazos extendidos hacia el frente, el cartón debe agarrarse por una esquina, extendiéndolo lo más alejado del cuerpo para que una de las caras sea vista por un sensor y la otra por el otro sensor, como se puede ver en la figura 3.8. También conviene decir que esta configuración se ve más beneficiada cuando los sensores no están directamente uno frente al otro, ya que puede causar interferencia con el otro sensor, y es recomendable también, como se ve en la figura 3.7, mantener los Kinects en alturas diferentes para evitar este posible problema.

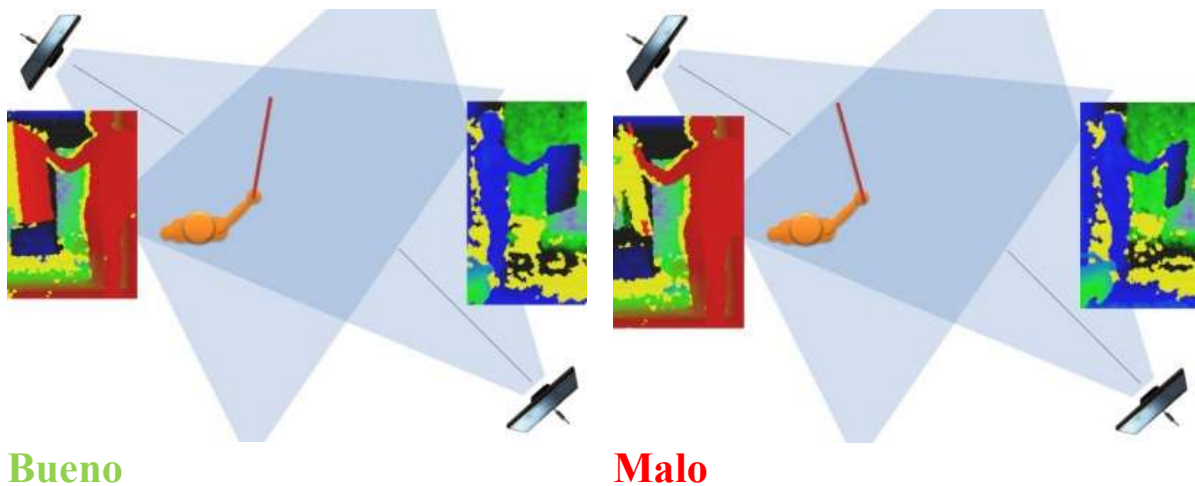


Fig. 3.8 Calibración del Kinect utilizando el software de iPiSoft con la segunda de las dos posibles configuraciones con 2 Kinect.

### 3.3.2. – Toma de prueba

La toma de prueba es la que se realiza para ver si la calibración ha sido exitosa, es relativamente corta y su propósito es asegurarse que la calibración haya sido la adecuada, se debe procesar para ver si sirvió o no. Si la calibración ha sido la adecuada, se procederá a grabar las tomas establecidas para el día.

### 3.3.3. – Tener una referencia tanto de audio como de video.

Se aconseja y es recomendable tener una referencia del audio como de video para cualquier tipo de sistema, ya que permite tener una referencia visual y de audio cuando se vaya hacer el procesamiento de los datos de captura, además puede ser utilizado como un place holder durante la edición. Esto debido a que de esta forma se puede avanzar en la postproducción, sin tener que depender o esperar el procesamiento de los datos que formarán parte de la escena final (Menache, 2010).

Conocer la posición de la cámara es importante también en todas las producciones audiovisuales, pero tiende a ser olvidada en la captura de movimiento, ya que no se está grabando la actuación final. Sin embargo, es importante tomarla en cuenta, porque cuando se captura la actuación siempre se quiere tener una referencia grabada. Si el proyecto es lineal, es mejor tener siempre al menos una cámara colocada en el ángulo y el encuadre final de la cámara, a objeto de ser utilizado como referencia al lado de la data capturada al momento de la post producción. La referencia con la grabada será utilizada en algunos casos para decidir entre las

tomas, o como un marcador de posición con fines de edición y definitivamente, como referencia a la hora de aplicar los datos al personaje final. Así mismo, es fundamental que el intérprete conozca dónde está la cámara cuando se actúa, lo cual viene a ser siempre donde está el público. Si no se sabe dónde está el público, no se puede establecer contacto con él a través del lenguaje corporal y la actuación se vuelve plana. Si no es posible colocar la cámara de vídeo en la dirección general de la cámara final, es necesario dejar que el actor sepa cuál es la posición final de la cámara.

### **3.3.5. – Organización de los archivos grabados.**

En los sistemas ópticos sin marcadores, es sumamente importante tener una buena organización de los archivos. Esto ayuda a tener claro a que captura pertenece cada archivo y así evitar que se confundan o extravíen. Por esta razón, es necesario establecer los planos de captura de movimiento propuestos por Menache (2010), ya que es uno de los métodos para la producción de captura de movimiento que facilita la organización y evita la confusión y pérdidas económicas en las grandes producciones, así como la frustración en las pequeñas producciones, motivado a no tener claridad a que archivo pertenece cada toma.

## **3.4. – Post-captura de movimiento**

Esta es la fase final del proceso de captura de movimiento. Aquí se hace el trabajo de edición, básicamente se termina y se pule el proyecto

### **3.4.1. – Revisar la data y mejorarla**

El software de IPiSoft tiene integrado varios procesos automatizados, los cuales facilitan la mejora y limpieza de la data capturada en el caso de que ésta sea necesaria, ya que algunas veces, a pesar de toda la preparación realizada pueden haber ciertas interferencias que hace que la data sea poco precisa o que requiera un tiempo de procesamiento extra. En la fase de limpieza y de mejoramiento de la data, se pueden presentar dos tipos de situaciones: la primera cuando los problemas pueden ser resueltos con el programa y la segunda, cuando los problemas requieren ser resueltos manualmente por el operador.

Las situaciones que pueden resolverse con el programa son dos: los tracking gaps y los cuadros individuales. Los tracking gaps básicamente ocurren cuando hay una oclusión, esto sucede principalmente con las manos, cuando una mano tapa a la otra, o cuando la pose es difícil

de distinguir, ejemplo cuando se cruza los brazos; hay oclusión porque la computadora le cuesta captar ese movimiento. Y la última, ocurre cuando el movimiento es muy rápido, lo que puede presentar errores en la captura de movimiento, porque al software se le dificulta anticipar la posición y pierde el rastreo. Para evitar esto, IPiSoft recomienda seguir el siguiente procedimiento:

- 1.- Se pasa dónde está el tracking gaps al próximo momento, donde la posición sea clara y se distinga fácilmente los movimientos.
- 2.- Si es necesario se utiliza las diferentes herramientas que provee el programa para que la pose del personaje sea igual a la del actor en la captura.
- 3.- Apagar el trajevtory filtering para que ésta no interfiere con las ediciones que se estén haciendo.
- 4.- Se le da al botón de Refit Pose para que la computadora entienda cuál es la pose que tiene que seguir el personaje.
- 5.- Activar el botón de track Backward, ya que esto hace que el programa comience a rastrear el movimiento, pero en reversa, lo cual permite arreglar el pedazo que tenía o presentaba el error.
- 6.- Parar el rastreo cuando ya se esté cerca del cuadro donde se encuentre el error, y de allí poco a poco se va haciendo el back tracking hasta llegar al último encuadre donde no ha tenido errores.
- 7.- Repetir el proceso hasta que sea necesario y se termine de realizar toda la toma.

En el caso de los cuadros individuales, se siguen los primero cuatro pasos anteriormente mencionados.

- 1.- Se busca el cuadro donde está el problema y se pasa al próximo momento donde la posición sea clara, y se distinga fácilmente los movimientos.
- 2.- Si es necesario utilice las diferentes herramientas que provee el programa para que la pose del personaje sea igual a la del actor en la captura.



- 3.- Apagar el Trajectory filtering para que éste no interfiera con las ediciones que se estén haciendo.
- 4.- Se le da al botón de Refit Pose para que la computadora entienda cuál es la pose que tiene que seguir el personaje.
- 5.- Se vuelve a activar el Trajectory filtering para guiar el programa.

Los problemas que no pueden resolverse utilizando el programa, son normalmente causados por la oclusión, para ello el operador puede tratar de guiar manualmente cuadro por cuadro la pose hasta que el programa pueda rastrearlo, si esto es muy problemático se recomienda entonces grabar de nuevo la toma y si se puede añadir sensores de profundidad extra mucho mejor, pero si esto no es posible, entonces se recomienda cambiar el movimiento por uno que sea más fácil de rastrear. Por esta razón, es que se hace necesario los ensayos para evitar este tipo de situación.

El programa tiene otras herramientas para ayudar a mejorar los movimientos capturados y hacerlos más suaves, que no se vean temblorosos. Antes se tenía que recurrir a otros programas especializados en el procesamiento de movimiento, pero IpiSoft lo ha integrado a este programa. Estas herramientas son las siguientes:

***Tracking refinement:*** que se basa en el algoritmo de rastreo del movimiento, pero es mucho más refinada y más intensivo, esta herramienta se debe reservar para el final, debido a que toma mucho tiempo de procesamiento, tres o cuatro veces más que el rastreo normal.

***Jitter Remove:*** es un filtro que se encarga de remover cualquier interferencia que pueda haber en la imagen, evita que ésta tiemble, hace que ésta sea más precisa y que los movimientos se vean más claros y naturales. Se recomienda en escenas de artes marciales, porque hace que los movimientos sean más precisos, claros e intimidantes.

***Trajectory filtering:*** funciona de una manera similar al anterior, lo que básicamente hace es limpiar el movimiento para quitar el temblor innecesario, pero no de una forma tan agresiva como la herramienta anterior. También puede utilizarse para pegar diferentes tomas que forman

parte de una secuencia de tomas combinadas, ya que facilita la transición de una toma a la otra sin que sea tan aparente.

### **3.4.2. – Aplicar la data al personaje 3D.**

Lo bueno de este programa es que está dirigido más a amateurs o pequeñas empresas que no pueden costear los equipos más tradicionales y sofisticados, aunque vale la pena mencionar que, varios estudios de cine y desarrolladores de video juego lo han utilizado en su proceso de producción. Debido a esto el programa debe ser capaz de salvar la data capturada en archivos de animación, en una gran variedad de formatos para que sea compatible con los programas básicos hasta los programas avanzados en el mercado con 3Dmax y Maya. Aquí se va a demostrar cómo se implementaría con el programa de Source Filmmaker, un programa gratuito, creado por la compañía VALVE para que la gente pueda crear cortos animados, utilizando los modelos de los personajes de sus videos juegos.

### **3.5. –Post-producción y Rendering.**

Esta es la fase final de cualquier producción audiovisual moderna, ya que todo se hace de manera digital. El rendering es como se termina la producción, dando así la película o corto final. En este trabajo no se profundizara en este aspecto, porque la captura de movimiento no afecta el desarrolla de esta fase, es un producto terminado al cual se aplicó la data capturada y pasa a ser parte de un proyecto de animación o de una producción convencional.

En esta fase se hace el trabajo de audio, edición, se aplican los efectos especiales, básicamente se termina y se pule el proyecto. Existe variedad de bibliografía que explica e ilustra este aspecto.

## Capítulo IV: Captura de Movimiento en la práctica

### 4.1.- Fase de preproducción

Como se explicó en el capítulo anterior, la principal razón de realizar la preproducción en la captura de movimiento obedece a la necesidad de reducir el costo y tiempo de producción, así como también prever y evitar problemas innecesarios. A continuación se presentan los pasos que se siguieron en la preproducción para la prueba de captura de movimiento.

#### 4.1.1.- Elaboración del guion

El guion es la idea elaborada, en el caso de este proyecto, se trata sobre una niña que juega y se imagina que está viajando por el espacio, luego de que su nave es estallada se despierta en un planeta habitado por diferentes criaturas, con las cuales termina bailando. Este guion se incluye en el anexo 2. Para su elaboración se siguió el formato estándar del Software Final Draft, el cual es considerado por muchos como el estándar de la industria hollywoodense, pero bajo los propósitos de esta prueba, se utilizó el programa de Microsoft Word 2013, ya que es más accesible, a pesar de no estar especializado para la elaboración de guiones cinematográficos, la figura 4.1 muestra el formato del guion.

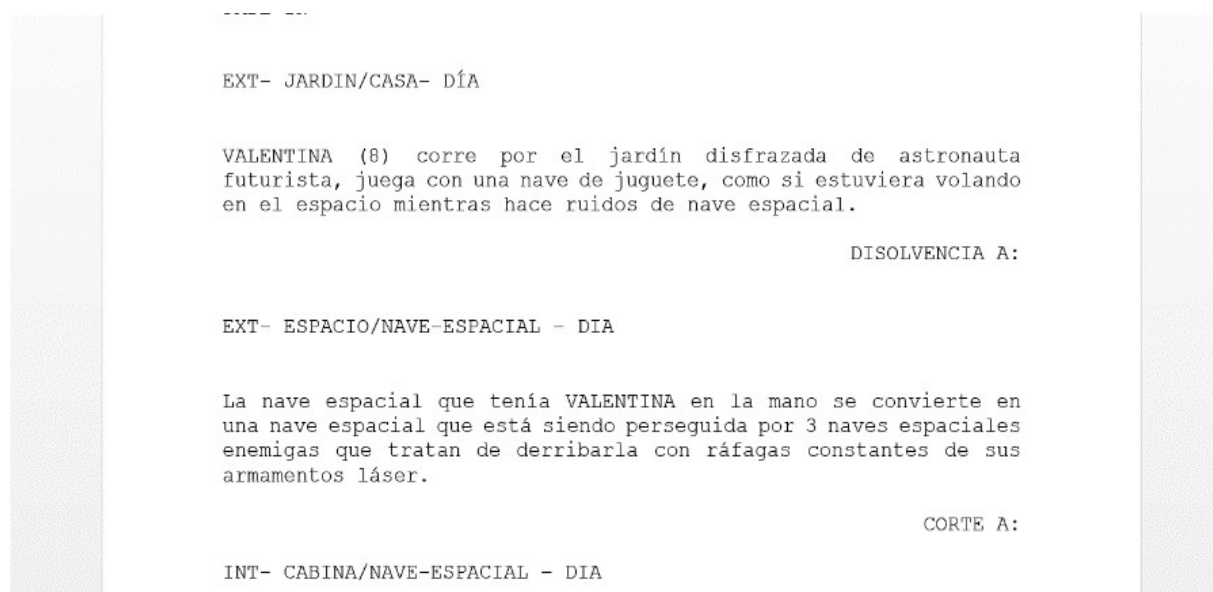


Fig. 4.1 Ejemplo de guión cinematográfico.

#### 4.1.2.- Pruebas de captura de movimiento

Según explica Menache (2011), este paso trata sobre la realización de pruebas de captura de movimiento, su propósito es verificar si el movimiento que se desea capturar es, en primer lugar, factible de aprehenderlo y en segundo lugar, cómo se ve animado antes de haber avanzado mucho en la producción. Esto es indispensable hacerlo, especialmente si se está utilizando por primera vez un sistema de captura; en el caso del presente trabajo se realizaron pruebas no solo de los movimientos que se deseaba utilizar en el corto de prueba, sino también de movimientos que pudieran ser complicadas de animar con métodos más tradicionales, a fin de poder así evaluar las virtudes de este sistema sin marcadores en ambiente de producción. Estos videos de captura forman parte del contenido incluido en los anexos digitales 1-3.

En la realización del presente trabajo se realizaron tres pruebas, una en la cual se utilizó un solo sensor Kinect, y dos donde se utilizaron 2 sensores Kinect en las dos posibles configuraciones, presentadas en las figuras. 4.2 y Fig. 4.3. El propósito de estas pruebas fue verificar si el sistema era capaz de realizar la captura de movimiento necesario para la ejecución de este estudio, y en el caso de constatar que no era capaz, tomar en cuenta estas limitaciones y así poder ajustar los movimientos, a fin de que funcionara dentro de las limitaciones del sistema.

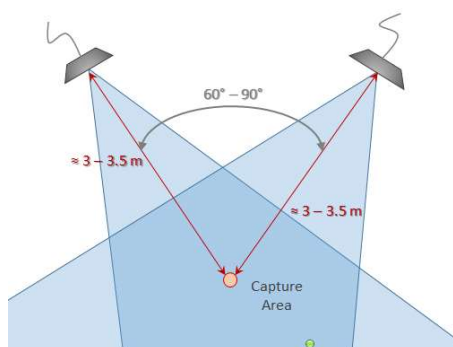


Fig. 4.2 Primera configuración con doble Kinect

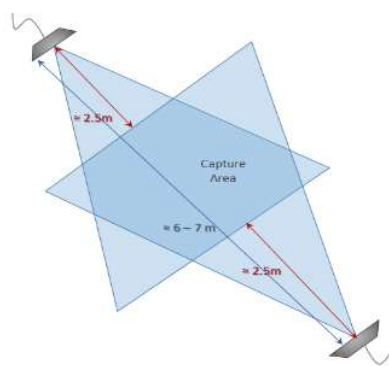


Fig. 4.3 Segunda configuración con doble Kinect

En la primera prueba se utilizó un solo sensor Kinect, debido a que es la forma más rápida y sencilla de las opciones disponibles del software de IPiSoft, el cual se puede realizar con una laptop y un sensor, sin necesidad de calibrar; además, es muy útil para hacer capturas de movimientos sencillas, especialmente donde hay movimientos frontales. Como se aprecia en la figura 4.4 es importante comenzar primero todos los movimientos, bien sea con uno o más

sensores, mediante una pose en T, que tenga una duración de cinco segundos, de modo de poder posicionar el esqueleto del software de iPiSoft con el actor que realizará los movimientos. Así mismo, en la Figura. 4.5, puede verse la imagen que representa la data capturada por el sensor de profundidad, y al lado aparece la versión grabada por la cámara RGB que forma parte del sensor Kinect.

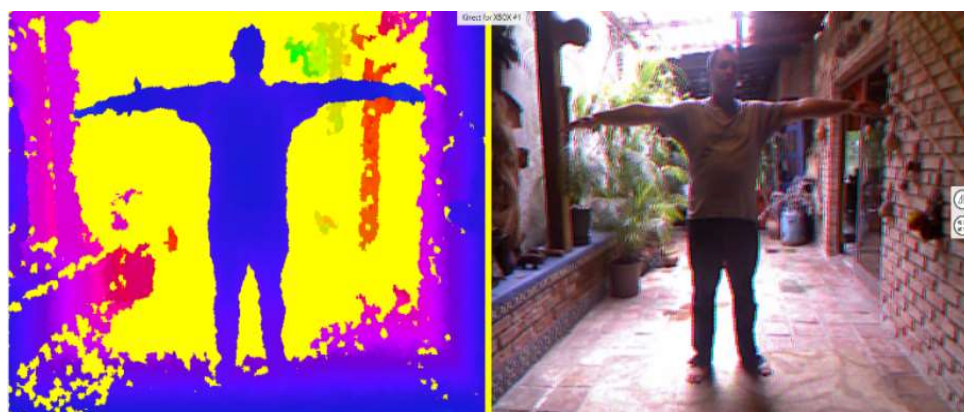


Fig. 4.4 Pose en "T", que es el pose para poder comenzar la captura

En la figura. 4.5, se ve el video ya procesado a través del software iPiSoft Mocap Studio, donde están superpuestos los datos de profundidad con la extracción del fondo, los datos de la cámara RGB, y el esqueleto propietario de iPiSoft con el cual se captura el movimiento, que luego se deben extrapolar a los huesos del esqueleto del modelo que se desea utilizar. En el caso del presente estudio, en las pruebas se trabajó con el formato .dmx para la animación y el formato .mdl para los modelos; porque estos son los únicos formatos compatibles con el programa de Source Filmmaker. Resulta esencial insistir que, es necesario probar y asegurarse que el esqueleto del modelo seleccionado sea compatible con el del software de iPiSoft, ya que si no lo es puede haber situaciones en que los esqueletos no cuadran y los movimientos no son compatibles, como se pueden apreciar en la Figura. 4.6, donde el esqueleto a la izquierda de color anaranjado es el esqueleto de iPiSoft y el de la derecha de color verde es el esqueleto del modelo, en el mismo se puede apreciar que no hay compatibilidad y se genera distorsiones en la posición de los huesos.

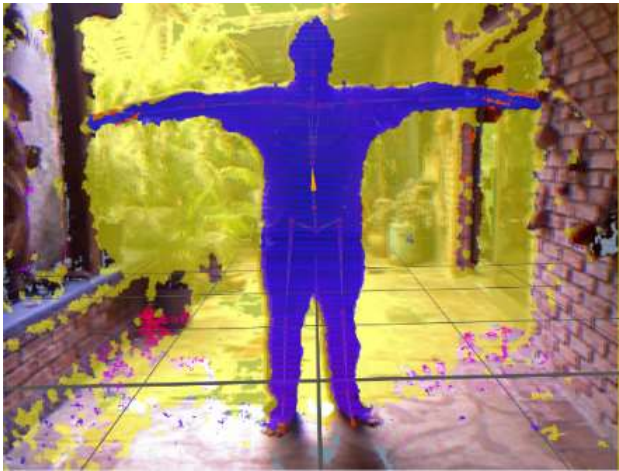


Fig.4.5 Pose en "T" cuadrada con la data de profundidad y el esqueleto de iPiSoft

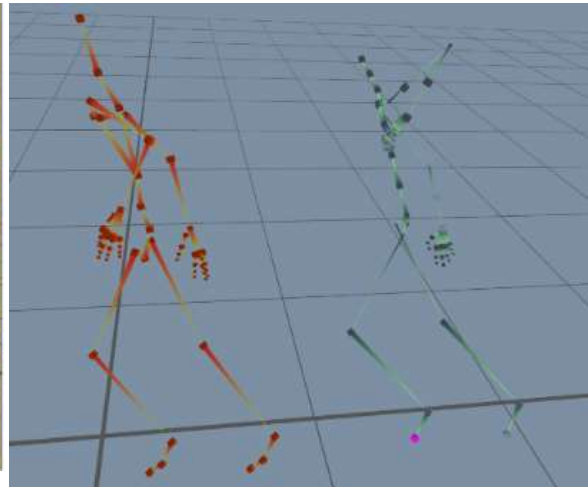


Fig.4.6 el esqueleto de la izquierda corresponde al esqueleto con el movimiento, mientras el de la derecha no coincide lo cual genera un error, lo que se evidencia con los brazos que no concuerdan.

El anexo digital 1 permite apreciar el nivel de distorsión que se puede llegar a tener. Esto es algo que siempre debe tomarse en cuenta, cuando se trabaja con modelos no personalizados. Ya que si los modelos son sacados de diferentes fuentes deben pasar por un proceso en el que se cambia el rigging o esqueleto del modelo y el formato de animación con el cual se trabaja, a fin de que éste sea compatible con iPiSoft y el software de animación escogido para el proyecto. Recuérdese que las pruebas se realizaron en el formato .dmx para la animación y el formato .mdl para los modelos; ya que estos son los únicos compatibles con el programa de Source Filmmaker.

#### Segunda prueba.

En la segunda prueba se utilizó la configuración de dos Kinect de manera similar a la presentada en la Fig.4.2, pero la cual si se tuvo que calibrar, a diferencia de la prueba anterior con un solo Kinect, cuya calibración es automática. La calibración de esta segunda prueba se realizó para ajustar la posición de ambos Kinect, y a objeto de que el programa conozca su posición en el espacio real y ajusten la posición en el espacio virtual, de modo de procesar la data capturada por los dos Kinect, no solo de forma individual sino de manera conjunta. El proceso de calibración se puede apreciar en la Fig. 4.7, que como se explicó en la capítulo anterior, se realiza con un cartón de dimensiones de 100cm de alto por 80cm de ancho; una vez que es calibrada la data, se guarda en un archivo aparte, a fin de utilizarla en todas las capturas que requiere esa calibración, ya que si se mueve uno de los Kinect todo el sistema se tiene que



recalibrar. Es importante señalar que, la configuración con dos Kinect permite capturar el doble de la data que la de la configuración con un solo Kinect, su uso se recomienda cuando haya movimientos frontales más exagerados y dinámicos, con rotaciones que puedan poner en occlusión alguna extremidad por tiempo prolongado, los cuales difícilmente pueden ser capturados por un solo sensor.



Fig. 4.7 Proceso de calibración de la primera configuración con doble Kinect



Fig. 4.8 Pose en "T" en la primera configuración con doble Kinect



Fig. 4.9 Prueba 2 de captura de movimiento en la primera configuración con doble Kinect

Otra de las pruebas realizadas, se hizo con un movimiento que causara errores, de modo de poder ejemplificar la forma cómo los mismos pueden aparecer en una sesión cotidiana de captura de movimiento, esto se aprecian en las Fig. 4.8 y Fig. 4.9, allí se evidencia la occlusión en ambos sensores Kinect, la cual causa graves errores, ya que no hay data que valide la posición del brazo derecho. El software al no tener los datos suficientes para poder posicionar el brazo, le da una animación aleatoria que no corresponde en ningún aspecto al movimiento real o al movimiento del resto del esqueleto del software.

En el caso de la prueba aquí realizada, se pudo solventar, utilizando una configuración diferente, la cual forma parte de la prueba 3; el motivo por el cual no se realizó la captura de estos movimientos, obedeció a que el espacio no permitía y la caminadora no se podía mover del lugar donde estaba ubicada, pero se aprovechó para mostrar los errores que pueden originarse en la captura, estos errores pueden evidenciarse más claramente en la Fig. 4.10 y la Fig. 4.11. Allí se puede ver que el brazo derecho no corresponde con la imagen y es una posición irreal. Esto se puede evidenciar con mayor claridad en el anexo digital 2.

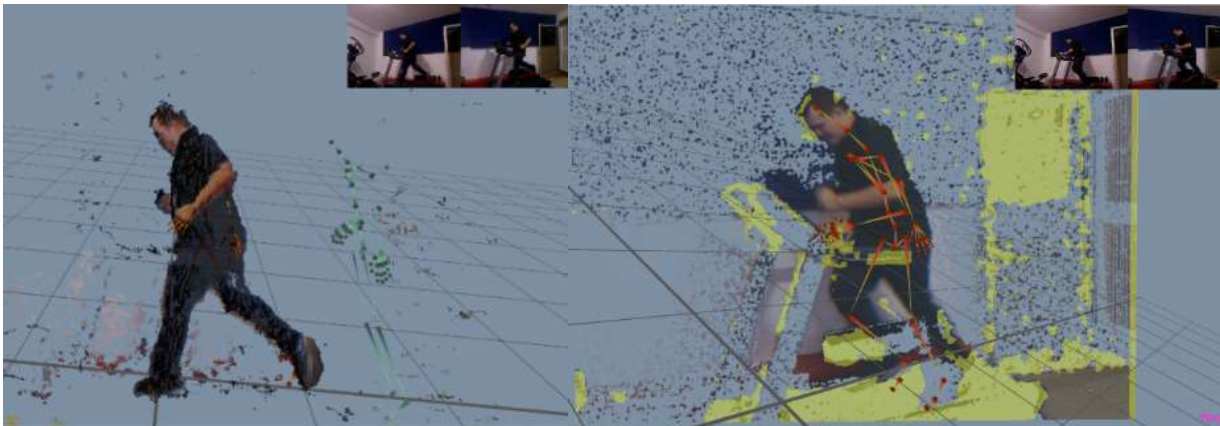


Fig.4.10. Vision de la camara 1 en donde se puede evidenciar que el movimiento del brazo derecho en el esqueleto no corresponde con el movimiento real, un error causado por la Oclusion.

Fig. 4.11. Vision de la camara 2 en donde se puede evidenciar que el movimiento del brazo derecho en el esqueleto no corresponde con el movimiento real, un error causado por la Oclusion.

### Tercera prueba.

La tercera prueba se realizó con la segunda configuración, la cual se presentó en la Fig. 4.3. Esta configuración resultó ser más problemática que las anteriores, debido a que se necesitaron 2 computadoras en lugar de una, motivado a que la extensión USB que se utilizó generaba errores esporádicos y hacía imposible concretar una sesión de captura.

Para solventar esta situación, se recurrió a interconectar más de una computadora y cada una de éstas se les conecta el respectivo sensor Kinect. Esto se logra conectando las distintas computadoras a la misma red inalámbrica o alámbrica, y dentro del programa de iPiSoft Recorder se establece cuál de las computadoras será la Maestra, que tiene la función de comenzar y parar las grabaciones, así como cuál(es) será(an) la(s) esclava(s) que seguirán el mando de la computadora Maestra. Este paso no siempre es necesario, si se está utilizando el sensor Kinect original, diseñado para el Xbox 360 y si se tiene una extensión USB estable, pero



es obligatorio cuando se utilizan múltiples sensores Kinect para el Xbox One, ya que los drivers son diferentes en ambos sensores y en el caso del Kinect para el Xbox One solo se puede tener un sensor conectado por computadora. También vale la pena decir, aunque no es relevante para la prueba aquí realizada, debido a que se utiliza el software de iPiSoft y éste no permite utilizar ambas versiones del Kinect de manera simultánea. Sin embargo, el software de Brekel Pro Face puede funcionar de manera simultánea al software de Brekel Pro Body, y pueden utilizarse de manera simultánea con la misma computadora un sensor Kinect Xbox One y un sensor Kinect Xbox 360.

Una vez conectadas las computadoras, desde la computadora maestra se inicia la sesión de grabación de manera similar a las sesiones anterior de prueba; es decir, que se graba una toma de calibración, pero a diferencia de la primera configuración el cartón se sostiene con una mano extendida alejado del cuerpo y se abanica lentamente acercándolo y alejándolo de los sensores, como se puede ver en la figura 4.12. Una vez que se grabó la toma de calibración y ésta se revisó, se pudieron grabar las tomas que se necesitaban de dicha configuración.



Fig. 4.12 Calibración de la tercera prueba, donde se evidencia la como se debe sostener el cartón con que se calibra el sistema.

Para procesar las tomas, sean éstas de calibración o las tomas con las acciones se deben pasar por un proceso de unificación, a objeto de tener toda la data que se guardó en cada una de las computadoras esclavas. Como fue el caso de esta prueba, se tuvo que importar la data de ambas computadoras en que se hicieron las grabaciones, los archivos que correspondían a cada toma, (2 archivos por toma) y se combinaron a través del programa iPiSoft Recorder, para así tener un solo archivo por cada toma y poder procesarlo con el programa de Mocap Studio. Al calibrar, el programa de iPiSoft Mocap Studio no siempre se detecta que las cámaras están una frente a la otra, si ese es el caso, se pueden utilizar las herramientas que ofrece el programa que

permite posicionarlas una frente a la otra, de modo que así el programa pueda determinar con mayor facilidad la posición en el espacio de cada cámara y la calibración sea válida.

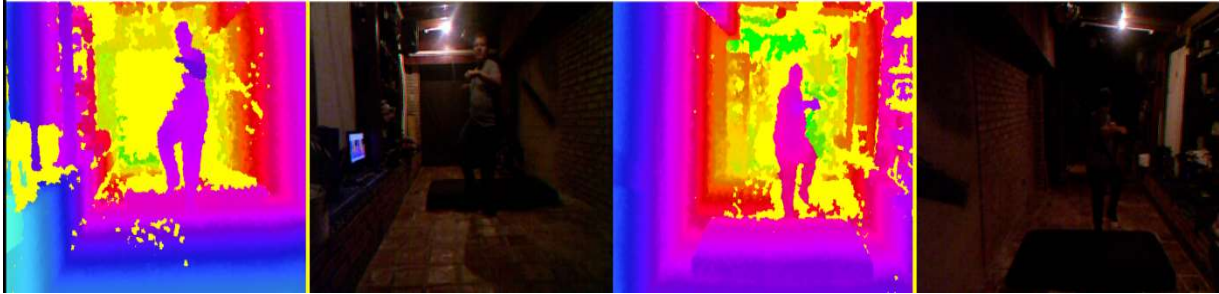


Fig. 4.13 Imágenes de la tercera prueba, en las que se puede evidenciar que los Kinect son capaces de rastrear el movimiento de ambos brazos a pesar de que estén solo visibles por uno de los dos sensores.

Como se puede ver en la figura 4.13 las cámaras tienen una visión total del cuerpo por ambos lados, de esta forma se minimizan las posibilidades de oclusión, y de las dos configuraciones disponibles con los dos Kinect, ésta es la que genera el volumen de trabajo más grande, además que proporcionan la mayor libertad de movimiento, ya que se pueden tener rotaciones del cuerpo en las que se daría la espalda a uno de los Kinect sin la preocupación de que cause errores en la captura. Esto también permite que el actor haga movimientos que son arduos para captar con la primera configuración de, no debido a la oclusión sino al límite de alcance óptimo del sensor, un ejemplo de esto es cuando el actor cae al piso con su cuerpo extendido como se puede ver en la figura 4.14. En el anexo digital 3, donde se muestra cómo se pueden capturar los diferentes movimientos que con las otras configuraciones hubieran sido imposible realizarse.

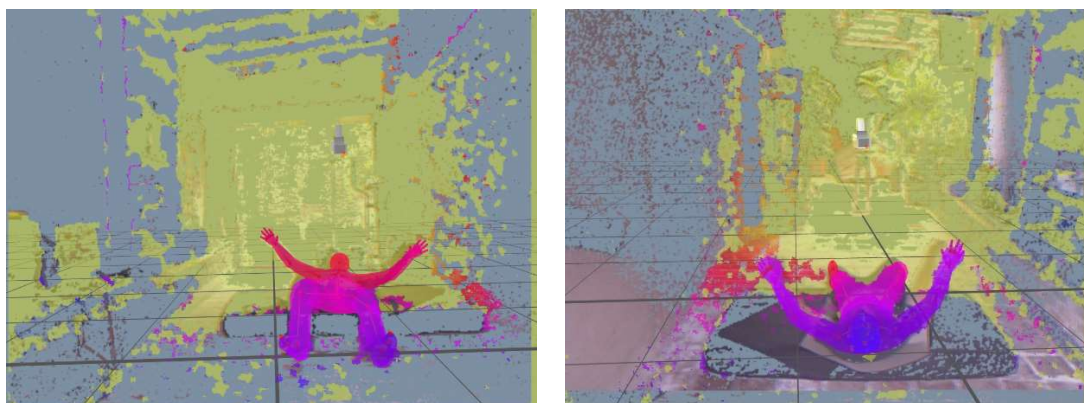


Fig. 4.14 Vistas de ambos sensores Kinect en las que el autor cae al piso.

A pesar de todas las ventajas, no se recomendaría el uso de esta configuración, excepto en aquellos casos que se necesite capturar movimientos específicos que no serían posibles con las otras configuraciones. Esto debido a la cantidad de espacio necesario para utilizar esta configuración y que requiere un proceso más largo para configurar y calibrar, lo cual puede causar retrasos durante una grabación. Pero si se dispone del espacio, especialmente, y si éste se puede dedicar como estudio de captura de movimiento y los equipos estén fijados en posiciones permanentes, ésta sería la configuración más favorable, aunque si este es el caso, se debería considerar el uso de cámaras Playstation Eye, ya que éstas tienen la ventajas sobre los sensores de profundidad, cuando el sistema se utiliza en un espacio dedicado y controlado, debido a que por cada sensor Kinect se pueden utilizar 3 cámaras Playstation Eye, y son capaces de tener un área de captura superior a la de los sensores de profundidad. Como se dijo en el capítulo 2, la principal ventaja que tienen los sensores de profundidad es que estos son más fácil de usar, transportar y configurar, lo cual permite que se pueda realizar la captura de movimiento en diferentes sitios, aunque se sacrifica un poco las dimensiones del área de captura.

Experiencia derivada de las pruebas.

Una vez cumplidas las pruebas de captura de movimiento, se derivaron varias conclusiones, no solo sobre el uso de los sistemas de captura de movimiento, sino sobre el uso de un sistema como el de iPiSoft, el cual se basa en la captura de movimiento sin marcadores. Entre esas conclusiones es imperativo señalar como primer punto que, uno debe ajustarse a la herramienta que dispone para poder aprender sus límites y sacarle el mejor provecho a la tecnología que se dispone. Esto es verdad no solo con la captura de movimiento, sino con todas las tecnologías que estén relacionadas al cine. Además, tiene una gran relación también con el hecho de que, mucha gente cree que la captura de movimiento es una solución mágica a los problemas de animación. Sin embargo, como se ha referido a lo largo de este trabajo, este no es el caso, no existen soluciones mágicas, especialmente con estos sistemas de bajo costo, que la característica que normalmente los hace sobre saltar del resto, es ese, su bajo costo. Esto significa que hay varias soluciones que son o más precisas o tienen un volumen de captura muy superior al de los sistemas de bajo costo.

El sistema aquí utilizado tiene un volumen limitado, y sirve para muchas acciones, como bailar o artes marciales, pero se tiene que tomar en cuenta que hay momentos que se puede salir

del área óptima de captura. Por ello, es recomendable que se marque cuáles son los límites del espacio después de que este esté establecido. A pesar de esto, si se puede decir que la captura de movimiento facilita el proceso de producción, cuando se necesita tener ciertas partes animadas en una producción audiovisual, ya que a pesar de sus limitaciones provee suficientes funciones que justifican su uso. Lo que sí resulta imperativo recomendar es tener un conocimiento avanzado sobre animación, a objeto de poder sacarle en mejor provecho posible al software. A pesar de ser un sistema lo suficientemente sencillo de operar para alguien que sin experiencia previa puede operarlo, tras unas pocas horas de práctica, también es cierto que el software se puede utilizar de manera adecuada; sin embargo, para poder utilizar la data capturada y sacarle su mayor potencial, se necesita tener un amplio conocimiento de los principios de la animación, mucho más allá de los tratados en este trabajo, porque se requiere saber utilizar programas de animación y edición de data de captura de movimiento. Y entre esos está el que se considera el estándar de la industria de la animación, como lo es la Suite de Autodesk, aunque se pueden utilizar también opciones libres, como Blender. Contar con este conocimiento facilita hacer composiciones de varios movimientos capturados, a objeto de que parezcan un solo movimiento fluido, así como modificar algunas imperfecciones que pudieron haber surgido en el proceso de captura.

Otra conclusión que debe referirse, tras las pruebas y que es similar a las dadas por Menache (2010), pero tras el proceso empírico de probar la tecnología es la siguiente: entre menos se muevan los sitios de grabación más eficientes son las sesiones de captura de data, a pesar de que sea rápido y considerablemente más sencillo, el proceso de posicionar y calibrar los sensores se puede realizar en apenas 20 minutos.

#### **4.1.3.- El plano de Captura de movimiento.**




A diferencia de las otras facetas de la preproducción, ésta es la que más se distingue de otros tipos de producciones audiovisuales, ya que según Menache (2010), está diseñada para facilitar el proceso de captura de movimiento, el cual puede ser complicado y confuso, especialmente cuando se está empezando en el área de la captura de movimiento y animación. En la figura 4.15 se puede apreciar el plano de captura de movimiento recomendado por este autor. En esta sección se ejemplificará la parte teórica que se trató en el capítulo anterior, se guiará en la elaboración de un plano de captura de movimiento, éste representó una de los

momentos más complejos en la preparación, ya que no tiene muchas relaciones con otros tipos de producciones audiovisuales. Como el plano de producción no se puede completar únicamente en la fase de preproducción, se especificara cuáles son las secciones del plano que se completan en los respectivos momentos de la producción hasta ser completado.

Al principio esto puede parecer abrumador, y puede serlo, si el proyecto no es lo suficientemente grande para requerirlo, como es el caso de las pruebas de este trabajo, en las cuales se utilizó una lista de tomas sobre este plano de captura, principalmente porque no se tenían más de 3 tomas planificadas, lo cual es suficientemente pequeño para que no haya muchas posibilidades de confusión, ya que se generan pocos archivos de captura, porque el trabajo no los requiere. Pero a medida que las producciones audiovisuales sean más grandes, en las cuales se requieran varias tomas de captura de movimiento, o en el caso que sea un proyecto de animación en el cual la gran mayoría se realice con captura de movimiento, entonces una lista de tomas sería insuficiente. Para evitar confusiones y mantener un buen orden, se puede decir que las complejas virtudes comunicacionales del plano salen a relucir, especialmente cuando se tienen diferentes grupos trabajando en captura y animación.

En la tabla 4.1 se explica cuál es el significado de los colores asignados al plano de captura de movimiento que se presenta como ejemplo, azul para identificar la sección que se llenaría en la fase de pre producción, salmón las que se llenarían antes de comenzar la grabación de la toma y verde cuales se llenarían al final de la sesión correspondiente a esa serie de tomas, esto se hizo con el fin de explicar de forma clara y precisa en que momentos de la producción se van completando las distinta secciones del plano. También se aprovechó recalcar algunas de las secciones que pudieron haber generado confusión en la misma figura 4.15, a pesar de que se ha tratado con profundidad en el capítulo anterior, se explicaran con brevedad para que sirva de guía en cómo utilizar esta herramienta que propone Menache.

Tabla 4.1

Color	Descripción
	Las secciones que estén coloreadas de este color corresponden a las secciones que se deben llenar durante la fase de pre-producción.
	Las secciones que estén coloreadas de este color corresponden a las secciones que se deben llenar durante la fase de producción, al comienzo de la sesión de captura de movimiento.
	Las secciones que estén coloreadas de este color corresponden a las secciones que se deben llenar durante la fase de producción, al final de la sesión de captura de movimiento.

Cada sección del plano de captura de movimiento indica qué información pertinente se debe colocar, con las excepciones de las secciones en las cuales el tamaño no lo permitía, lo cual se explicará a continuación.

- FPS Final: son los cuadros por segundo que se van a utilizar en el proyecto.
- FPS Puros: son los cuadros por segundos con los cual funciona el sistema de captura de movimiento.
- Unidad de medida: es la unidad que se va utilizar en el proyecto (Ej. cm)
- Conversión: es lo que equivale una unidad digital en correspondencia a la unidad de medida.
- Eje Global: en el plano el eje global se trabaja a través de las vistas (frontal, aérea y lateral) lo cual permiten establecer el eje cartesiano.
- Esqueleto: hay 2 secciones para el esqueleto, una donde se anota el nombre de éste, y la otra donde se pone una referencia gráfica.
- Modelo: imagen de los modelos que se vayan a utilizar en la toma.

En la sección donde se enlistan los personajes y los actores que los representan, se debe indicar si estarán haciendo captura de manera simultánea o de manera individual.

En esta sección se ponen los planos que deben ser realizados en la toma o tomas combinadas que formen parte de las tomas				Descripción:	Una pequeña descripción de lo que pasa en la toma	Nombre Archivo:	Asignación														
				Dialogo:	El dialogo que corresponde a la toma de captura	Esqueleto:	Nombre del esqueleto														
<p>Lista de personajes</p> <table border="1"> <tr> <td>1</td> <td>Valentina</td> <td>1</td> <td></td> </tr> <tr> <td>2</td> <td></td> <td></td> <td></td> </tr> <tr> <td>3</td> <td></td> <td></td> <td></td> </tr> </table>	1	Valentina	1		2				3						<p>Vista delantera</p>	<p>Vista Aérea</p>	<p>Vista de perfil</p>	<p>Esqueleto:</p>	<p>Modelo:</p>	<p>Volumen Total: 1.5m³</p>	<p>Interacciones</p>
	1	Valentina	1																		
	2																				
3																					
<p>Calibración</p> <p>Nombre de la calibración correspondiente a la toma</p>																					
<p>Marcadores</p> <p>No aplica ya que no hay marcadores</p>																					
FPS Final:		FPS Puros:		Cuadros finales: La cantidad total de cuadros en la toma																	
Comienzo: en que momento de la toma comienza el movimiento		Fin: en que momento de la toma termina el movimiento		Dur:		Duración total de la toma															
Handles:		Duración del handle		Duración total:		Duración total del archivo															
Unidad de medida		Conversion																			
<p>Instrucciones especiales: En esta sección se puede colocar cualquier instrucción especial que no se puede expresar en otra sección del plano</p>																					
				Tramoya y utilería		Tomas a ser compuestas															
				1 En esta sección se enlistará cualquier pieza de tramoya o de utilería que se necesita para poder completar la escena.		Nombre															
				2		Nombres de las tomas que deben ser combinadas con la actual toma.															
				3		Desde que punto debe comenzar a ser compuesta la toma referida															
				4																	
				5																	
				6																	
				7																	
				7																	

Fig. 4.15 Plano de Captura de Movimiento como ejemplo de como realizarlo.



## 4.2.- Fase de producción

En esta sección trataremos de como fue el proceso de producción, enfocándonos en el proceso de captura, aunque se mencionara como fue el proceso de grabación tradicional con los actores en vivo, pero solo los aspectos que incumben a las necesidades de grabación que afectarían la inclusión de los efectos especiales animados, así como la transferencia de los movimientos a los actores digitales. En esta prueba no se recurrió a la creación de modelos digitales, debido a que ello sobrepasaba la intención y objetivos del trabajo. En su lugar, se utilizaron modelos digitales que se encuentran de manera gratuita a través del Steam Workshop, Garry's Mod y la página web de gamemaps.com en su sección para el juego *Left 4 Dead*, cuyo uso es permitido de forma gratuita, siempre y cuando sean utilizados sin fines de lucro, La selección de estos modelos se hizo motivado a que son compatibles para funcionar con el programa de animación Source Filmmaker, y que su rigging o esqueleto sea compatible con el software de iPiSoft. Además se tomó en cuenta que a pesar de tener una figura humanoide fueran modelos diversos. .

### 4.2.1 Producción Tradicional

La fase de producción, donde se grabó de manera tradicional no fue divergente de los métodos que forman parte de una producción, de los cuales se habló en el capítulo tres La única acotación especial a realizar es que las tomas fueron grabadas en vivo, se grabaron en 30 cuadros por segundo, porque ello es similar a la velocidad de 30 Hz del sensor Kinect que grabó los movimientos. Ello se hizo para evitar cualquier posible desfasamiento del movimiento en relación con el video durante la fase edición, a objeto de evitar diferentes velocidades de cuadro por segundo.

La producción tradicional se realizó en dos días. El primer día se trabajó con la niña actriz, para ello fue necesario realizar varios ensayos para las diferentes tomas, tanto con la cámara, como sin ella, debido a la inexperiencia previa de la actriz, a fin de acostumbrarla a no ver a la cámara. El segundo día de trabajo se dedicó a grabar las voces para las partes animadas, de modo de poder medir los tiempos que deben durar la toma animada que formaban parte del proceso de captura.



### 4.2.2 Producción de Captura de Movimiento

La captura de movimiento, ya en la fase de producción no se diferencia mucho de las pruebas que se realizaron en la fase de pre producción, porque se utilizaron las 3 configuraciones que se probaron, la mayor diferencia estuvo en que se aprovechó los conocimientos adquiridos durante dichas pruebas, principalmente la distorsión causada por la luz solar, que abrumaba a los sensores infrarrojos del Kinect, lo cual generaba fallas en la data y dificultaba la calibración. Es necesario tomar en cuenta que esta etapa en la fase de producción, se trata de generar la data de movimiento necesaria, a fin de que se pueda realizar la animación que requiere el corto. Se debe seguir la lista de tomas y el plano de captura de movimiento que se hizo previo a la etapa de pre-producción, la misma consiste en una toma con un solo Kinect, una toma con dos Kinect en la primera configuración (Fig. 4.2) y tomas con la segunda configuración (Fig. 4.3) con 2 sensores Kinect.

La mayor diferencia entre una sesión de captura de movimiento de prueba, con una sesión de producción es el refinamiento que pasa la data después de ser capturada, a fin de asegurar que la data de movimiento disponible para la transferencia sea la mejor que se pudo obtener de la toma seleccionada, esto se condensa en que la data de pruebas solo tiene un proceso mínimo para solo revisar las capacidades del sistema y asegurarse que las configuraciones eran válidas, mientras que el proceso de producción final exige el uso de varios filtros de movimientos, al igual que el proceso de refinación del procesamiento de la data. Esto significa que, cada escena se tardara mucho más tiempo en procesar, pero el resultado será el de más alta calidad.

Una vez capturada y procesada toda la data, esta información debe transferirse al software de animación que se haya decidido utilizar, como ya se ha dicho en capítulos anteriores. En la demostración que se realizó en este trabajo, se utilizó el Source Filmmaker, no solo por su fácil uso, y por ser diseñado para ser utilizados por entusiastas y amateurs en la animación, sino porque se basa en tecnología de videojuegos, que es la temática central de este trabajo. El proceso específico de transferencia será abordado más adelante en este capítulo.

### 4.2.3 Proceso de Animación

El proceso de animación se realizó con el software de Source Filmmaker, el cual es de uso gratuito y funciona como un híbrido con video juego, porque utiliza el motor de videojuegos Source y un programa de animación que permite realizar cortos animados, utilizando recursos de diferentes videojuegos que emplean el mismo motor. En comparación con otros programas de animación, éste está más limitado en sus capacidades, así como también se ve más limitado en la calidad de la imagen final, pero tiene la ventaja de que es más accesible que otro software, sumado a ello requiere menos conocimiento técnico y cuenta con una biblioteca de objetos propios relativamente extensa; todo lo cual facilita que principiantes en el mundo de la animación lo utilicen para crear sus propios cortos animados. El software cuenta además con muchas de las mismas herramientas y controles que los programas de animación profesional poseen, pero es capaz también de realizar animación de pose a pose.

Los modelos utilizados durante la fase de animación se seleccionaron durante la etapa de preproducción y fueron confirmados durante las pruebas de captura de movimiento por ser compatibles con el esqueleto del software de iPiSoft. Los tres modelos utilizados se pueden ver en las figuras 4.15.



Fig. 4.15. Modelos Utilizados en la prueba del trabajo.

Para importar la data del software de iPiSoft Mocap Studio, se necesitó utilizar el archivo propietario de Valve para la aplicación de data de animación .dmx. El cual en la actualidad se puede importar de manera directa al programa Source Filmmaker en cuatro sencillos pasos, gracias a una de las actualizaciones del software de iPiSoft, debido a la popularidad del Source Filmmaker, generada por los entusiastas de la animación, quienes son los principales

compradores de este sistema de captura de movimiento de bajo costo. Hay otros métodos para poder importar la data de la animación al Source Filmmaker, pero este requiere del uso de un software extra, cuyo uso no es gratuito, exige de una licencia para estudiantes que proveen sus creadores, el Motion Builder de Autodesk. Este es un software que se especializa en la creación de movimientos animados, bien sea esto desde cero, o la aplicación de data de distintas fuentes, entre ellas la data de software de captura de movimiento. Dicho método provee la ventaja de tener más control sobre la data de animación y crear con mayor facilidad movimientos compuestos por varias tomas de captura de movimiento, cuando se compara con Source Filmmaker. Sin embargo, para poder utilizar el software se requiere un nivel de conocimiento que va más allá de la animación básica, y el programa no tiene el mismo nivel intuitivo que Source, porque fue diseñado como un programa introductorio a la animación. Por estas razones, en este trabajo se enfocó en el empleo del método más directo que no requiere el uso de Software extra.

En concreto, aquí se realizó lo siguiente, en primer lugar fue necesario importar el personaje (o su esqueleto) a iPi Mocap Studio para la transferencia de movimiento. Esto se hizo mediante la importación de un archivo DMX en su pose predeterminada; es decir sin ninguna animación. Este archivo posee la información sobre el esqueleto del modelo, lo cual es suficiente para la transferencia del movimiento. A objeto de crear el archivo, se procedió de la siguiente manera:

1. Se buscó Animation Set Editor Tab,
2. Se aplicó el clic en el botón con el símbolo de + y luego clic en **Create Animation Set for New Model.**
3. Se eligió el modelo que se deseaba utilizar y se aplicó clic en **Open.**
4. Al efectuar clic derecho sobre el modelo y seleccionar **Export Animation, se abrió la ventana de exportación de animación.** Hay una casilla llamada ASCII, la cual se debe seleccionar.
5. Se le dio el nombre correspondiente y se escogió la carpeta en que se desea salvar la animación, hecho esto se le dio al botón **Save.**

La animación exportada está en el formato DMX con la información del modelo en ASCII, el archivo puede ser importado a iPiSoft Mocap Studio, para poder realizar la transferencia de los movimientos, se debe cumplir lo siguiente:

1. Buscar el Export Tab.
2. Hacer clic en **Target Character** y seleccionar la opción **Import from file** o simplemente se presionan las teclas **Ctrl+T** para abrir la ventana de manera directa.
3. En la ventana que se abrió, se abre el archivo DMX que se desea y se hace clic en **Open**.
4. Si el esqueleto del modelo es el estándar que desarrollo Valve, los huesos del esqueleto son asignados automáticamente, si no es el caso, cada hueso debe ser asignado con su homólogo en el esqueleto de iPiSoft, a objeto de poder transferir el movimiento. Esto se hace de la siguiente forma:
  - Se selecciona el hueso que se desea asignar de la lista en el panel derecho.
  - Se le hace clic al botón con la imagen de un Cursor.
  - Por último se selecciona el hueso correspondiente en el esqueleto que fue importado.
  - Si se desea, se puede guardar la relación con el esqueleto para que no se tenga que repetir este proceso.
5. Al estar asignados los huesos al esqueleto importado, se exporta la animación, haciendo clic en el botón de **Export Animation**.

El archivo creado puede ser importado a Source Filmmaker para realizar la animación, de este punto en adelante el proceso de la creación de las escenas es similar a la creación de un proyecto audiovisual animado.

### 4.3.- Fase de Post-producción

Durante la fase de Postproducción, la cual es similar en la mayoría de los casos, independientemente de si el proyecto es realizado con animación o grabaciones más tradicionales, la única excepción es el añadir los efectos especiales visuales, ya que los mismos son normalmente hechos con imágenes generadas por computadora, y son incluidos durante la fase de producción en producciones animadas. Se debe tener presente que, la postproducción

no es un solo proceso único, sino un grupo de procesos que ocurren después de las grabaciones que forman parte de la fase de producción. La fase de post-producción en la creación de una película, por lo general, toma más tiempo que el rodaje, y puede tomar varios meses en completarse, especialmente en grandes producciones, ya que incluye varios procesos que tienden a ser:

- La edición del video; en el caso del presente trabajo se realizó con el software de Adobe Premiere y Adobe After Effects para generar las transiciones especializadas.
- La creación o edición del soundtrack del proyecto.
- Foley: especialmente en proyectos animados, donde no se genera una base de audio como en los proyectos tradicionales.
- La edición de sonido, lo que incluye el balance de los diálogos, el añadir efectos de sonidos.
- La corrección de color, lo cual normalmente se hace con software especializado en las grandes producciones. En el caso de este proyecto, se utilizó el programa de Adobe After Effects.
- La última fase de toda postproducción es la renderización final del proyecto.

El proceso de edición de una película también es visto como una segunda dirección, porque a través de la post-producción es posible cambiar la intención de la película. De la misma forma que el tono de una película es determinada por el director con la selección del encuadre de las escenas y su cinematografía. Mediante la edición, con el uso de la corrección de color, la adición de música y sonido, se puede cambiar la atmósfera general de la producción, e influenciar fuertemente el cómo la audiencia percibe la película.



## Conclusiones

La animación tiene su génesis en los inicios de la humanidad, existe desde antes del cine. Animar es dar vida a objetos, de ahí que los primeros hombres se inspiraron con la sombra y crearon animación cuando danzaban a la luz del fuego. Romaguerra y Ramio (1999, p. 45) precisa que animar es una manera genuina de crear movimiento mediante unas imágenes realizadas con una técnica, tan específica como fotografiar uno a uno. El arribo de la animación al mundo artístico ocurre en el teatro con el praxinoscopio de Reynaud en 1888, el cual permitía ofrecer un espectáculo a partir de la proyección de dibujos animados móviles, pero su repunte se ensombreció en 1895 con la entrada de la primera exhibición cinematográfica, realizada por los hermanos Lumière, al punto que los dibujos animados no retornaron sino hasta 10 años más tarde.

A partir de 1907 surgieron distintos tipos de animaciones, entre ellas la animación cuadro por cuadro, que se volvió popular. El film *Fantasmagoría* producido en 1908 es un ejemplo de esto y se le reconoció como la primera película animada; con ella se impuso la idea de calcar y luego fotografiar los dibujos, a fin de evitar saltos en la imagen y movimientos bruscos, imponiéndose lo que hoy se conoce como animación tradicional, llamada también animación dibujada a mano, que fue el proceso más utilizado para hacer películas animadas durante el siglo XX.

Un primer antecedente de lo que se conoce como captura de movimiento, se dio con los trabajos realizados por Muybridge en 1860, quien logró tomar una foto para comprobar que las cuatro patas de un caballo estaban en el aire al mismo tiempo en un punto del galope; ello lo estimuló a trabajar en la idea de capturar el movimiento de un animal, y para lograrlo empleó una serie de cámaras que se iban activando automáticamente con el galopar del caballo, lo que le permitió perfeccionar la técnica. Su técnica fue luego reconocida académicamente y se publicaron sus trabajos.

Otro predecesor de la captura de movimiento es Marey, fisiólogo francés, quien desarrolló la cronofotografía, inspirado por los trabajos de Muybridge, pero descartó el uso de múltiples cámaras, por considerarlo poco científico. Así en 1882 logra desarrollar la pistola

cronofotográfica, la cual se apoya en una rueda para tomar rápidamente una secuencia de fotos, de 12 cuadros por segundo. También diseñó unos trajes especiales, de color negro y material refractante, con los cuales se podía capturar el movimiento humano, sin apreciarse el cuerpo del mismo.

Fleischer es considerado otro predecesor de la captura de movimiento, ya que creó el Rotoscopio en 1917, sustentado en la idea de que la animación se podía hacer con mayor facilidad y con movimientos más naturales. Su método sustenta la idea de que los actores realizan las escenas que se desean animar, a fin de procesarlas y revelarlas, el film se coloca en la máquina de rotoscopio y se proyecta a la mesa de trabajo donde el animador «calca» sobre cada fotograma, para así transmitir al dibujo la naturalidad y secuencialidad de movimientos. Disney tomó y utilizó luego esta herramienta con el fin de estudiar el movimiento y no para la animación, porque la consideraba poco artística.

Lo innegable es que el cine dentro del campo de las artes es el que más se vincula con los avances tecnológicos, desde sus inicios fue considerado una innovación tecnológica. Los progresos tecnológicos que ha experimentado el cine son cada vez más diversos y sofisticados, los mismos han contribuido significativamente en la producción de obras cinematográficas. En la actualidad las Imágenes Generadas por Computadora (CGI) juegan un papel trascendental, comparables con el paso del cine mudo al sonoro, del cine blanco y negro al de color. Al punto que la tecnología ha cambiado el modo en que las películas se producen, en especial con la llegada del cine digital, ya que constituye una herramienta al alcance del artista, que le amplió el campo de sus posibilidades para la creación cinematográfica.

La captura de movimiento en el campo del entretenimiento es el descendiente del rotoscopio, pero éste se centraba en la animación en dos dimensiones (2D), porque fue diseñado para los dibujos animados tradicionales. Con el surgimiento de la animación en tres dimensiones (3D) se generó el nacimiento de una nueva forma de rotoscopio para 3D, que es en esencia la captura de movimiento, ya que busca la automatización del proceso mediante el empleo de la computadora, la cual es capaz de rastrear el movimiento. Esta tecnología fue impulsada a finales de la década de los 70 y principio de la década de los 80 del siglo pasado por iniciativas de las



universidades norteamericanas, enfocada en la captura de movimiento óptico; pero su proyección pública se realizó en el área de la publicidad a mediados de los 80.

Los grandes avances tecnológicos en la captura de movimiento se dieron en el año 1988, entre ellos figura la captura de movimiento electromagnética, que permitía su realización en tiempo real, y fue usada en un programa de televisión. La industria del cine fue incursionando progresivamente en la captura de movimiento, se crea un dispositivo mecánico denominado "exoesqueleto". La primera película en la que un personaje digital se creó con éxito, utilizando la captura de movimiento, fue Toys. La empresa PDI fabricó una serie de accesorios, destinados a mantener los sensores en su lugar y hacer más fácil la experiencia de capturar movimiento en todos los proyectos relacionados con el exoesqueleto electromagnético.

Los equipos para la captura de movimiento en la década del 90 estaban destinados al área de la medicina y el ejército. La mayoría de sistemas de captura de movimientos que se ofrecían eran ópticos. Luego emergieron compañías especializadas en dar este servicio y sus principales clientes eran los de la industria del videojuego, porque sus consolas se volvieron más poderosas y necesitaban que los movimientos fueran más realistas. Ya para comienzos del siglo XXI los sistemas de captura de movimiento se tornaron más accesibles, tanto en costo como en capacidad técnica, porque las empresas buscaron ampliar su mercado, introduciendo sistemas de animación accesible para la captura de movimiento. Lo que estimuló el surgimiento de diversas compañías que ofrecían servicios de captura de movimiento, y a su vez las grandes productoras de videojuegos crearan su propia división de captura de movimiento para ofrecer servicio desde la misma compañía sin intermediarios.

La captura de movimiento es una forma de traducir los movimientos de un "actor" en vivo, a una representación digital que mueve a un "actor digital", se puede utilizar para grabar el movimiento de cualquier sujeto que tenga movimiento propio y aplicarlo a su homólogo digital. Existen diversos mecanismos y sistemas para la captura de movimiento, todos ellos se basan en los mismos principios. Se trata de rastrear los puntos claves del sujeto, los cuales representan las diferentes áreas de movimiento del sujeto, que no son más que los puntos de pivotes que conectan las partes rígidas del sujeto. En el caso del ser humano, los puntos pivotes son las articulaciones.

Existe variedad de tecnologías disponibles para la captura de movimiento como los sistemas ópticos y mecánicos. Cada tipo tiene sus ventajas y desventajas, dependiendo del área en la cual se aplique. Se pueden clasificar en Sistemas ópticos y no ópticos. **Los ópticos** recurren a la simple utilización de data capturada a través de sensores de imagen, a fin de triangular la posición del sujeto en un espacio 3D, relativo a la posición de dos o más cámaras que hayan sido previamente calibradas para proporcionar proyecciones superpuestas. Se clasifican en sistemas con marcadores pasivos y activos, así como en sistemas sin marcadores; estos últimos pueden ser: con múltiples cámaras o con sensores de profundidad. **Los Sistemas no ópticos** son mucho más variados que los ópticos, algunos de ellos funcionan bajo sus mismos principios, mientras que otros utilizan una variedad de sensores como potenciómetros y acelerómetros para capturar el movimiento.

La captura del movimiento humano, se ha realizado con una variedad de tecnologías. Cada día hay más avances, entre estos está la captura de movimiento sin marcadores, esta innovación se desarrolló a bajo costo en la industria de los videojuegos, la cual hizo accesible la captura de movimiento en los hogares, a través de las consolas de los videos juego. Un ejemplo de ello es el *Xbox Kinect*, con el cual se rompió el paradigma de emplear un control para poder interactuar con los videojuegos. Este equipo posee un sensor de profundidad tridimensional, tiene la capacidad de captar el movimiento del usuario en tiempo real y a la precisión suficiente para que el jugador interactúe con el mundo virtual que ve en la pantalla de su televisor. Esta tecnología funciona con un sensor infrarrojo que crea una nueva forma de capturar data, además es capaz de grabar en tres dimensiones. Sin embargo, el Xbox Kinect no revolucionó el mundo de los videojuegos como sus creadores esperaban, pero fue acogido por la comunidad mundial de inventores e innovadores, en particular por los departamentos de tecnología de varias universidades que han desarrollado proyectos en torno a la capacidad de captura de movimiento y sensor de profundidad.

Uno de los programas que facilitó la captura de movimiento a través del uso del Kinect, es el desarrollado por Nikonov, a través de su compañía **iPiSoft**. Su intención fue hacer más accesible a las masas el uso de la captura de movimiento y logró un sistema óptico de captura de movimiento más práctico. El sistema de iPiSoft es óptico sin marcadores, consiste en dos programas diferentes, el iPi Recorder y el IPi Mocap Studio. Este sistema está diseñado para

procesar data en forma de imagen, ya que la tarea principal de iPi Mocap Studio es procesar las imágenes del video y reconocer el patrón de la figura humana para así rastrear su movimiento a lo largo del video. De este modo, el usuario tiene la posibilidad de escoger las porciones del video que desean procesar y seleccionar la sección de video que más le interesa.

Para la captura de movimiento es necesario disponer, tanto de un hardware con alta capacidad como de un software adecuado. Al igual requiere de una minuciosa planificación, que permita tener una idea clara sobre qué es lo que debe hacerse para poder lograr un proceso satisfactorio y un mejor producto. Por consiguiente, la captura de movimiento hace indispensable cumplir con estos cinco grandes pasos: la preproducción, el trabajo de captura de movimiento, la captura de movimiento, la postcaptura de movimiento, la postproducción y la renderización. A medida que la tecnología es más avanzada la fase de captura se vuelve más fácil de realizar, porque la tecnología se encarga de procesar la información capturada. Esos cinco pasos se siguieron en la prueba que se ejecutó en el presente estudio.

Es importante tener presente que el incursionar en las tecnologías de video juegos para la captura de movimiento, es básico ser modesto y ajustarse a las herramientas que se dispone, si se tienen tecnologías sofisticadas los resultados serán óptimos, pero si no se dispone de esa tecnología hay que aceptar que se confrontan limitaciones; de ahí que lo primero que debe aceptar quien se inicia en estas tareas, es ajustarse a la herramienta que dispone para poder aprender sus limitaciones y sacarle el mejor provecho a la tecnología con la cual cuenta. Esto último no es solamente valido para la captura de movimiento, sino para todos los procesos que se tienen que cumplir al recurrir al uso de tecnologías en la producción cinematográfica.

Es cierto que en la actualidad se cuenta con variadas soluciones para la captura de movimiento, tanto de alto como de bajo costo. Las primeras tienen un volumen de captura muy superior al de los sistemas de bajo costo. Sin embargo, el sistema que en el presente estudio se utilizó, siendo de bajo costo y tener un volumen limitado demostró que sirve para capturar movimientos variados como: caminar, correr, saltar, bailar. Con su uso se comprobó que se puede emplear en la animación de cortos. Además, se puso en evidencia que es accesibles para fomentar la animación en la producción de cortos, tanto por su costo como las posibilidades que ofrece. Para su uso es aconsejable que se marque bien los límites del espacio donde se va

realizar la captura, a fin de evitar salirse del área óptima de captura. Así mismo, como lo planteaba Menache entre menos se muevan las locaciones más eficientes son las sesiones de captura de data. Por igual, a objeto de evitar errores en la captura de movimiento, se recomienda utilizar ropas que no sean tan holgadas, es preferible el empleo de ropa más ajustada que facilite la definición de la silueta humana, de colores oscuros para evitar las sombras, especialmente en casos donde la iluminación no es uniforme.

En esa misma dirección hay que recalcar que la captura de movimiento no es una solución mágica para los problemas de animación, constituye una herramienta que requiere conocimiento y pericia, la cual solo se gana con el entrenamiento y la práctica, para iniciarse en la tarea de capturar movimiento es necesario seguir las sugerencias de los pioneros en el tema, como se realizó en este trabajo. El poner a prueba la captura de movimiento recurriendo al Kinect y el software de IPISoft solo me hizo evidente la necesidad de contar con profundos y amplios conocimientos en animación, lo cual va mucho más allá de los meros principios de la animación, de ahí que es indispensable seguir profundizando en el tema.

Hoy se puede afirmar que la captura de movimiento es el futuro de la masificación de la animación, ello en parte obedece a que estamos en una sociedad eminentemente visual, las relaciones se dan a través de imágenes, y éstas viajan, recorren grandes distancias, somos casi testigos presenciales de eventos y manifestaciones sociales, culturales y naturales que ocurren en distintas regiones del mundo sin movernos del mueble en que estamos sentados o acostados en nuestros hogares. Es una realidad que esta tecnología de sensores de profundidad posibilitan la captura de movimiento sin marcadores de manera sencilla y rápida, no requieren procesos complicados de calibración y es más tolerable a diferentes condiciones de iluminación; además se está incluyendo ya en los Smartphone, incluso su difusión aparece como anuncio este año en la última exposición del CES 2016, en la cual Google anuncio que con el Proyecto Tango esperan traer visión computarizada, sensores de profundidad y capacidades para el rastreo de movimiento en los celulares de Lenovo para finales de este año. Así mismo, Compañías como Apple están buscando integrar la tecnología PrimeSense a sus celulares, desarrollada por la misma compañía que impulso las tecnología detrás del Kinect, y por su parte Intel está buscando también promover con las manufactureras de celulares Smartphone sus sensores de

profundidad, Realsense (Grunewald, 2016). Por consiguiente, estas tecnologías tendrán un gran impacto en el futuro y en la forma como nosotros interactuaremos con el mundo que nos rodea.

El cine aprovechará y sacará la máxima utilidad de todos estos progresos, al igual que lo ha venido haciendo en el devenir de su desarrollo. De hecho, hoy se crece la probabilidad de que estas tecnologías de captura de movimiento, al llegar a los celulares genere un boom creativo, al igual que ocurrió cuando en los últimos años se logró el desarrollo de las cámaras DSLR y conquistaron la capacidad para grabar video, con lo cual se abrieron las puertas a la gente creativa, que emergió con cortos y proyectos audiovisuales, trayendo a su vez la democratización de este tipo de tecnologías y técnicas, las cuales estaban antes reservadas solo para los profesionales del área.



## Glosario

**Animación:** es crear la ilusión del movimiento a una serie de dibujos e imágenes. Se considera como una ilusión óptica que dota de vida, movimiento a objetos que por su propia cuenta no lo tienen.

**Animación Keyframe:** es un método de animación en lo cual se enfoca en los cuadros o frames claves de una animación los cuales son hechos por el animador principal en una animación tradicional o por el animador en una animación digital, y los cuadros entre los diferentes cuadros clave o keyframe son hechos por los asistentes del animador o en el caso de la animación digital son hechos por la computadora.

**Adobe Premiere:** software para la edición de video basado en una línea de tiempo, la versión más actualizada es conocida formalmente como Adobe Premiere Pro y forma parte del Adobe Creative Cloud.

**Adobe After Effects:** es un software que tiene un formato de estudio destinado para la creación o aplicación en una composición, así como realización de gráficos profesionales en movimiento y efectos especiales, que desde sus raíces han consistido básicamente en la superposición de capas. Adobe After Effects es uno de los softwares basados en línea de tiempo más potentes del mercado. Al igual que Adobe Premiere Pro, forma parte del Adobe Creative Cloud.

**AutoDesk:** es una compañía dedicada al software de diseño en 2D y 3D para las industrias de manufacturas, infraestructuras, construcción, medios y entretenimiento y datos transmitidos vía inalámbrica. Autodesk fue fundada en 1982.

**Autodesk Motion Builder:** es un software de animación de personajes 3D profesional producido por Autodesk. Se utiliza para la producción virtual, la captura de movimiento, y la animación de keyframe tradicional.

**Blender:** programa de uso gratuito de multi plataforma, dedicado especialmente al modelado, iluminación, renderizado, animación y creación de gráficos tridimensionales. También de composición digital que utiliza la técnica procesal de nodos, edición de vídeo, escultura (incluye topología dinámica) y pintura digital. En Blender, además, se puede desarrollar video juegos ya que posee un motor de juegos interno.

**Captura de movimiento:** proceso mediante el cual se puede grabar el movimiento en vivo y traducirlo a datos matemáticos que puedan ser utilizados para rastrear el movimiento de puntos claves en un espacio tridimensional en un periodo de tiempo determinado. Es traducir los movimientos de un “actor” en vivo, a una representación digital que mueve a un “actor digital”.

**CGI:** siglas en inglés de Imágenes Generadas por Computadora

**CCD:** (charge- coupled device). Son un sensor de cámaras digitales sensibles a la luz que utilizan una matriz de células fotoeléctricas (también llamados píxeles) para captar la luz, y luego medir la intensidad de la luz en cada una de las células y así crear una representación digital de la imagen.

**CMOS:** (Complementary metal–oxide–semiconductor) es una de las familias lógicas empleadas en la fabricación de circuitos integrados. En la fotografía digital y la captura de video se utiliza comúnmente para hablar de los sensores de píxeles activos, o APS por sus siglas en inglés. Los cuales son sensores de imágenes con circuitos integrados CMOS, que han surgido desde el 2010 como una alternativa más económica y fácil de manufacturar a los sensores CCD.

**Código de tiempo (time code):** es una referencia precisa del tiempo, que nos ayuda a sincronizar y permite la compatibilidad del intercambio de material. Más concretamente es una señal digital electrónica de alta frecuencia que consiste en un conjunto de pulsos producidos por un generador de códigos de tiempo, y que se graba a lo largo de la longitud de un canal de audio, o en la pista cue de una cinta de video, de la misma manera que se graba la pista de sonido convencional.

**Dispositivo plug-n-play:** es un dispositivo que sigue los protocolos PnP (Plug and Play). Estos protocolos facilitan el descubrimiento de componentes de hardware por un sistema sin la necesidad de configurar de manera física el dispositivo, o que no requieran que el usuario intervenga cuando se presenten conflictos de recursos, en la actualidad la mayoría de los dispositivos digitales son plug and play.

**Fenómeno *phi*:** es la ilusión óptica que ocurre debido a la persistencia retiniana que lleva a nuestro cerebro a percibir un movimiento continuo en donde hay una sucesión de imágenes.

**iPiSoft:** compañía rusa que desarrolla un sistema para la captura de movimiento accesible a las masas o personas que se están iniciando en el mundo de la animación.

**iPi Recorder:** programa para capturar, reproducir los registros de vídeo de múltiples cámaras o sensores de profundidad. Es decir su propósito es grabar el video.

**iPi Mocap Studio:** programa que analiza el registro y traduce el movimiento físico del actor en data que se le puede aplicar a uno o varios actores virtuales, es decir su propósito es procesar el video.

**Kinect:** es un accesorio de videojuego con sensor de profundidad, que permite realizar captura de movimiento de baja fidelidad en tiempo real, y de mediana a alta fidelidad cuando es procesada por una computadora.

**Mapa de Textura:** El mapeo de textura originalmente se refería a un método (que ahora se llama con más precisión diffuse mapping) que simplemente se envuelve y se asigna píxeles de una textura a una superficie 3D. En las últimas décadas el advenimiento de las tecnologías de renderización de pasos múltiples, El mapeo de texturas se ha vuelto más complejos tal como: mapeo de altura (height mapping), bump mapping, mapeo normal, mapeo de desplazamiento (displacement mapping), mapeo de la reflexión reflection (mapping), mapeo especular (specular mapping), mipmaps, mapeo de oclusión (occlusion mapping) y muchas otras variaciones en la técnica (controlado por un sistema de materiales). Estos avances han hecho posible simular con realismo casi fotográfico en tiempo real, mediante la reducción en gran medida del número de polígonos y de cálculos de iluminación necesarios para construir una escena 3D realista y funcional.



**Megapíxeles:** es un millón de píxeles, normalmente utilizado para expresar el número de elementos que posee un sensor de imagen. El número de píxeles puede ser el total que tiene el sensor o los efectivos, es decir los que funcionan de manera óptima.

**Modelos 3D:** están compuestos por polígonos, con planos de colores sólidos, a los cuales se le puede aplicar un mapa de textura.

**Motion-controllers:** controles a base de movimiento, estos normalmente son sistemas híbridos de captura de movimiento, los cuales se basan en sistemas ópticos y en sensores como acelerómetros para determinar el movimiento.

**Motion Capture Blueprint:** es un plano de captura de movimiento, mezcla entre el guión gráfico y una lista de tomas como los que se utilizan en proyectos cinematográficos convencionales, pero con información adicional de la captura de movimiento.

**Ray-tracing o trazado de rayos:** es un método de renderización basado en un algoritmo para síntesis de imágenes tridimensionales. Propuesto inicialmente por Turner Whitted en 1980, está basado en el algoritmo de determinación de superficies visibles de Arthur Appel denominado Ray Casting (1968)

**Renderizar:** es un anglicanismo de la palabra render, que literalmente significa presentar, pero en la computación se refiere al proceso de generar una imagen o vídeo mediante el cálculo de iluminación GI partiendo de un modelo en 3D. Este término técnico es utilizado por los animadores o productores audiovisuales (CG) y en programas de diseño en 3D.

**Sensores de Profundidad:** son sensores de imágenes avanzados que son capaces de percibir profundidad en la imagen, la mayoría son monocromáticos y necesitan un sensor de imágenes tradicional para poder asignar color a la data capturada.

**Sistemas ópticos:** son aquellos capaces de capturar movimiento a través de sensores de imagen, pueden ser con marcadores o sin marcadores. Se basan en los mismos principios matemáticos de la triangulación, que utiliza la trigonometría para determinar posiciones de puntos o medidas de distancia.

**Sistemas ópticos con marcadores:** sistema de captura de movimiento que usa marcadores especiales, colocados a un actor. Pueden ser con marcadores pasivos y activos.

**Sistemas ópticos sin marcadores:** sistema de captura de movimiento con visión artificial o inteligencia computarizada, capaz de reconocer la silueta humana, rastrearla a lo largo de los cuadros y segmentarla en un esqueleto digital.

**SMPTE:** es un protocolo de código de tiempo desarrollado a principios de 1967, la Society of Motion Picture and Television Engineers (SMPTE), propuso a los distintos fabricantes crear un único código de tiempo estándar, y así evitar problemas en la compatibilidad

**Vicon 8:** primer sistema óptico digital diseñado con gráficos por computadora, su software introdujo la grabación de videos en el formato AVI, al mismo tiempo, de la captura de movimiento, es capaz de soportar código de tiempo SMPTE. Desarrollado por la compañía Vicon

**Wire-frame:** es como un modelo 3D se ve como cuando los mapas e incluso las caras poligonales se han eliminado para dejar sólo los contornos de sus polígonos que lo componen, esto consiste en puntos vectoriales conectados por líneas. Un alambre también se puede llamar una malla de alambre.

**Workstations:** Son computadoras especializadas, que se diferencian de las computadoras personales, ya que están diseñadas para cumplir ciertas labores técnicas, sean estas en aplicaciones científicas o en aplicaciones profesionales como la animación o edición de video.

**Xbox Kinect** equipo de video juego diseñado para capturar movimiento en tiempo real y que funciona como un dispositivo plug-n-play y sin necesidad de marcadores

**Cámaras RGB:** son cámaras a color que utilizan sensores de imágenes tradicionales, sean estos CCD o CMOS

.

## Bibliografía

- About me. (2010, August 21). Retrieved from <http://brekel.com/about/>
- Adamson, A., & Jenson, V. (2001). *Shrek*.
- Alsina, H., & Romaguera, J. (2007). *Textos y Manifiestos del Cine: Estética. Escuelas. Movimientos. Disciplinas. Innovaciones* (Edición: edición). Madrid: Cátedra.
- Auzel, D., & Tchernia, P. (2000). *Emile Reynaud et l'image s'anima*. Paris: Editions du May.
- Avlastimova, I. (2010, March 16). Interview with director iPi Soft - Mikhail Nikonov [Pagina Web]. Retrieved from [http://www.render.ru/books/show\\_book.php?book\\_id=877#](http://www.render.ru/books/show_book.php?book_id=877#)
- Azéma, M., & Rivère, F. (2012). Animation in Palaeolithic art: a pre-echo of cinema. *Antiquity*, 86(332), 316–324.
- Borenstein, G. (2012). *Making Things See: 3D vision with Kinect, Processing, Arduino, and MakerBot* (1 edition). Sebastopol, CA: Maker Media, Inc.
- Borshukov, G., Piponi, D., Larsen, O., Lewis, J. P., & Tempelaar-Lietz, C. (2005). Universal Capture - Image-based Facial Animation for “The Matrix Reloaded.” In *ACM SIGGRAPH 2005 Courses*. New York, NY, USA: ACM. <http://doi.org/10.1145/1198555.1198596>
- Braun, M. (1995). *Picturing Time: The Work of Etienne-Jules Marey* (Reprint edition). Chicago: University Of Chicago Press.
- Brekel Pro Face. (2012, July 1). Retrieved from <http://brekel.com/brekel-kinect-pro-face/>
- Brekel Pro Face 2. (2015, March 13). Retrieved from <http://brekel.com/brekel-pro-face-2/>
- Brown, B. (2011). *Cinematography: Theory and Practice: Image Making for Cinematographers and Directors* (2 edition). Amsterdam ; Boston: Focal Press.
- Cameron, J. (1989). *The Abyss*.

Cameron, J. (1991). *Terminator 2: Judgment Day*.

Charles Csuri, (1968). *Hummingbird (1968)*. Retrieved from  
<https://www.youtube.com/watch?v=awvQp1TdBqc>

Clegg, B. (2007). *The Man Who Stopped Time: The Illuminating Story of Eadweard Muybridge -- Pioneer Photographer, Father of the Motion Picture, Murderer* (2nd Printing edition).  
Washington, D.C: Joseph Henry Press.

Cohl, É. (1908). *A Fantasy*.

Cottrell, W., Hand, D., Jackson, W., Morey, L., Pearce, P., & Sharpsteen, B. (1938). *Snow White and the Seven Dwarfs*.

Courtet-Cohl, P., Génin, B., & Takahata, I. (2008). *Émile Cohl : L'inventeur du dessin animé. Avec 2 dvd-rom* (1re édition). Sophia-Antipolis: Omniscience.

Crichton, M. (1973). *Westworld*.

Fleischer, D. (1918). *Out of the Inkwell*.

Fleischer, M. (1917, October 9). Method of producing moving-picture cartoons. Retrieved from  
<http://www.google.com/patents/US1242674>

Glaze, V. (2009). *100 Años de cine*. (C. Fujiwara, Ed.). Barcelona: BLUME.

Glebas, F. (2008). *Directing the Story: Professional Storytelling and Storyboarding Techniques for Live Action and Animation* (1 edition). Amsterdam ; Boston: Focal Press.

Grunewald, S. J. (2016, February 2). Depth-Sensing Cameras Will Soon Turn Every Smartphone into a High-Quality 3D Scanner. Retrieved from  
<https://3dprint.com/117809/depth-sensing-phone-cameras/>

Hessler, G. (1974). *The Golden Voyage of Sinbad*.

- Huls, A. (2013, April 4). The Jurassic Park Period: How CGI Dinosaurs Transformed Film Forever. *The Atlantic*. Retrieved from <http://www.theatlantic.com/entertainment/archive/2013/04/the-i-jurassic-park-i-period-how-cgi-dinosaurs-transformed-film-forever/274669/>
- Ilyin, Y. (2010, May 19). The harsh Russian Motion Capture. Retrieved March 26, 2016, from <http://old.computerra.ru/terralab/multimedia/532492/>
- Johnston, O., & Thomas, F. (1995). *The Illusion of Life: Disney Animation* (Rev Sub edition). New York: Disney Editions.
- Khoshelham, K., & Elberink, S. O. (2012). Accuracy and Resolution of Kinect Depth Data for Indoor Mapping Applications. *Sensors*, 12(12), 1437–1454. <http://doi.org/10.3390/s120201437>
- Kitagawa, M., & Windsor, B. (2008). *MoCap for Artists: Workflow and Techniques for Motion Capture* (PAP/CDR edition). Amsterdam ; Boston: Focal Press.
- Klette, R. (2014). *Concise Computer Vision: An Introduction into Theory and Algorithms*. Springer Science & Business Media.
- Kuchera, B. (2010, September 1). Enough to forget the Wii? A week with the PlayStation Move. Retrieved June 27, 2016, from <http://arstechnica.com/gaming/news/2010/09/playstation-move-1.ars>
- Lasseter, J. (1995). *Toy Story*.
- Lisberger, S. (1982). *TRON*.
- Lucas, G. (1977). *Star Wars: Episode IV - A New Hope*.

- Marks, R. (2001). *Enhanced Reality: A New Frontier for Computer Entertainment*. Presented at the SIGGRAPH 2001, Los Angeles, California. Retrieved from [http://www.research.scea.com/research/pdfs/SIGGRAPHsketch2001RICK\\_pp.pdf](http://www.research.scea.com/research/pdfs/SIGGRAPHsketch2001RICK_pp.pdf)
- Melgar, E. R., & Diez, C. C. (2012). *Arduino and Kinect Projects: Design, Build, Blow Their Minds* (2012 edition). New York, N.Y. : New York: Apress.
- Menache, A. (2010). *Understanding Motion Capture for Computer Animation, Second Edition* (2 edition). Burlington, MA: Morgan Kaufmann.
- MITCSAIL. (2010). *Kinect Hand Detection*. Retrieved from [https://www.youtube.com/watch?v=tlLschoMhuE&feature=player\\_embedded](https://www.youtube.com/watch?v=tlLschoMhuE&feature=player_embedded)
- Nowell-Smith, G. (Ed.). (1999). *The Oxford History of World Cinema* (17th edition). Princeton, N.J.: Oxford University Press.
- Peter Foldes. (1974). *La Faim (Hunger)*. Retrieved from <https://www.youtube.com/watch?v=vwU3UARE6yc>
- PLAYSTATION EYE - Q+A | Three Speech: Semi-Official PlayStation Blog. (n.d.). Retrieved from <http://threespeech.com/blog/2007/04/playstation-eye-qa/>
- Price, D. A. (2009). *The Pixar Touch* (Vintage edition). New York: Vintage.
- Pro Body 2 – FAQ. (2014, October 24). Retrieved from <http://brekel.com/brekel-pro-body-v2/pro-body-2-faq/>
- Pro Body – FAQ. (2012, December 27). Retrieved from <http://brekel.com/brekel-kinect-pro-body/faq/>
- Ramió, J. R. i. (1999). *El Lenguaje Cinematográfico: Gramática, Géneros, Estilos y Materiales*. Ediciones de la Torre.
- Reeves, M. (2014). *Dawn of the Planet of the Apes*.

- Regazzoni, D., de Vecchi, G., & Rizzi, C. (2014). RGB cams vs RGB-D sensors: Low cost motion capture technologies performances and limitations. *Journal of Manufacturing Systems*, 33(4), 719–728. <http://doi.org/10.1016/j.jmsy.2014.07.011>
- Saltzman, M. (2010, September 21). Review: How Move is better than the Wii - CNN.com [News Outlet]. Retrieved June 27, 2016, from <http://edition.cnn.com/2010/TECH/gaming.gadgets/09/21/sony.move.review/>
- Scott, R. (1979). *Alien*.
- Sommers, S. (2004). *Van Helsing*.
- Sony Eye specs. (n.d.). Retrieved June 16, 2016, from <https://www.engadget.com/products/sony/playstation/eye/specs/>
- Sony PlayStation Eye Camera Specs. (n.d.). Retrieved June 27, 2016, from <http://www.cnet.com/products/sony-playstation-eye-camera/specs/>
- Spielberg, S. (1989). *Indiana Jones and the Last Crusade*.
- Spielberg, S. (1993). *Jurassic Park*.
- Spielberg, S. (1994). *Schindler's List*.
- Sterling, B. (2011, May 10). Augmented Reality: Kinect fitting-room for TopShop, Moscow. Retrieved June 18, 2016, from <http://www.wired.com/2011/05/augmented-reality-kinect-fitting-room-for-topshop-moscow/>
- Tobon, R. (2010). *The Mocap Book: A Practical Guide to the Art of Motion Capture*. (A. Restrepo, Ed.) (1st edition). Orlando, Fla.: Foris Force.
- TRONG DUONG, S. (2012). *INTERACTIVE FULL-BODY MOTION CAPTURE USING INFRARED SENSOR NETWORK*. University of Colorado.
- Trousdale, G., & Wise, K. (1991). *Beauty and the Beast*.

- University of Minnesota researchers studying the use of robots and computer vision to diagnose mental disorders in children. (2011, February 16). Retrieved June 18, 2016, from <http://discover.umn.edu/news/science-technology/university-minnesota-researchers-studying-use-robots-and-computer-vision>
- Veis, G. (1963). Optical tracking of artificial satellites. *Space Science Reviews*, 2(2), 250–296. <http://doi.org/10.1007/BF00216781>
- Visual Effects on Terminator 2 | Animator Mag - Archive. (n.d.). Retrieved from <http://www.animatormag.com/archive/issue-30/issue-30-page-14/>
- Wachowski, L., & Wachowski, L. (2003a). *The Matrix Reloaded*.
- Wachowski, L., & Wachowski, L. (2003b). *The Matrix Revolutions*.
- Walker, R. (2012, May 31). How Kinect Spawned a Commercial Ecosystem. *The New York Times*. Retrieved from <http://www.nytimes.com/2012/06/03/magazine/how-kinect-spawned-a-commercial-ecosystem.html>
- Watercutter, A. (2012, May 24). 35 Years After Star Wars, Effects Whiz Phil Tippett Is Slowly Crafting a Mad God. Retrieved July 18, 2016, from <http://www.wired.com/2012/05/phil-tippett-feature/>
- Weinberger, M., Sep. 8, 2015, 73, 093, & 23. (n.d.). The downfall of Kinect: Why Microsoft gave up on its most promising product. Retrieved June 16, 2016, from <http://www.businessinsider.com/why-microsoft-xbox-kinect-didnt-take-off-2015-9>
- Xbox On. (n.d.). *Kinect for Xbox 360 - You Are The Controller*. Retrieved from <https://www.youtube.com/watch?v=V32dnwrwjfQ>
- Yonemoto, S., Matsumoto, A., Arita, D., & Taniguchi, R. I. (1999). A real-time motion capture system with multiple camera fusion. In *International Conference on Image Analysis and*



*Processing*, 1999. *Proceedings* (pp. 600–605).

<http://doi.org/10.1109/ICIAP.1999.797662>

Zemeckis, R. (2004). *The Polar Express*.



## **Anexos**

**Anexo 1**

FADE-IN

EXT- JARDIN/CASA- DÍA

VALENTINA (8) corre por el jardín disfrazada de astronauta futurista, juega con una nave de juguete, como si estuviera volando en el espacio mientras hace ruidos de nave espacial.

DISOLVENCIA A:

EXT- ESPACIO/NAVE-ESPACIAL - DIA

La nave espacial que tenía VALENTINA en la mano se convierte en una nave espacial que está siendo perseguida por 3 naves espaciales enemigas que tratan de derribarla con ráfagas constantes de sus armamentos láser.

CORTE A:

INT- CABINA/NAVE-ESPACIAL - DIA

Una luz intermitente roja ilumina la cabina mientras VALENTINA, con un traje espacial oscuro con detalles en rojo y blanco, pilotea agresivamente la nave, con movimientos drásticos y exagerados, mientras trata de comunicarse por la radio.

VALENTINA

(Nerviosa

Ayuda, los rebeldes están siguiéndome,  
necesito ayuda

CORTE A:

EXT- ESPACIO/NAVE-ESPACIAL - DIA

La nave espacial de VALENTINA, está esquivando el ataque enemigo, cuando

VALENTINA

¡Me Dieron! ¡Me Dieron! ¡Me Dieron!

CORTE A:

EXT- JARDIN/CASA- DÍA

VALENTINA cae al piso y suelta su nave espacial de juguete.

DISOLVENCIA A:

EXT- PLANETA ALIENIGENA/SELVA- NOCHE

Una creatura se mueve por la selva, y observa a VALENTINA mientras se levanta, y se preparándose para abalanzarse sobre ella.

VALENTINA

Ahora estoy en medio de una selva.

La creatura se abalanza sobre VALENTINA

VALENTINA  
(Asustada Grita)  
AAAAAAHHHHHHHH

VALENTINA y el Alien se asustan

VALENTINA  
Ahora bailamos

VALENTINA baila con la creatura, y comienzan a aparecer otras

## Anexo 2. Kinect SDK Release Notes

### What's New

*Kinect for Windows 1.5, 1.6, 1.7, 1.8*

This section provides release notes for each released version of the Kinect for Windows SDK and the Developer Toolkit.

- [What's New in version 1.8 of the SDK and the Developer Toolkit](#)
- [What's New in version 1.7 of the SDK and the Developer Toolkit](#)
- [What's New in version 1.6 of the SDK and the Developer Toolkit](#)
- [What's New in version 1.5.2 of the Developer Toolkit](#)
- [What's New in version 1.5.1 of the Developer Toolkit](#)
- [What's New in version 1.5 of the SDK and the Developer Toolkit](#)

### What's New in version 1.8 of the SDK and the Developer Toolkit


Here's a link to the [known issues](#) for this release.

This release provides many new features, all part of the Developer Toolkit 1.8.0. The SDK/Runtime v1.8 will contain minor changes.


### Kinect Background Removal

The new Background Removal API provides "green screen" capabilities for a single person. The user can be specified using skeleton ID. The BackgroundRemovedColorStream API uses various image processing techniques to improve the stability and accuracy of the player mask originally contained in each depth frame. The stream can be configured to select any single player as the foreground and remove the remaining color pixels from the scene.

- New Background Removal samples demonstrate basic use of the Background Removal APIs.

 <b>Note</b>
Previous Green Screen samples in 1.7 have been renamed to CoordinateMappingBasics in 1.8.

- Native DLLs for Background Removal are KinectBackgroundRemoval180\_32.dll and KinectBackgroundRemoval180\_64.dll under %KINECT\_TOOLKIT\_DIR%\Redist. Managed DLLs for Background Removal are "Microsoft.Kinect.Toolkit.BackgroundRemoval.dll" under %KINECT\_TOOLKIT\_DIR%\Redist.

 <b>Note</b>
Use of managed DLLs requires use of native DLLs.

### Webserver for Kinect Data Streams

New web components and samples give HTML5 applications access to Kinect data for interactions and visualization. This is intended to allow HTML5 applications running in a browser to connect to the sensor through a server running on the local machine. Use this to create kiosk applications

on dedicated machines. The webserver component is a template that can be used as-is or modified as needed.

#### Note

The Webserver components and sample require .NET 4.5 (Visual Studio 2012) and Windows 8 (or later) for web socket functionality. The Kinect for Windows SDK JavaScript APIs support Internet Explorer 10 and later, Mozilla Firefox, and Google Chrome.

Microsoft.Samples.Kinect.Webserver is a webserver component that provides:

- Web socket end points that expose Kinect interactions (hand pointer movement plus push and grip gestures), background removal, user viewer, skeleton data, etc. as data streams and events sent from server to a HTML5-capable browser client
- REST end point that support GET operations to retrieve current stream configuration and POST operations to modify current stream configuration
- REST end point that acts as a simple file server that serves static content

WebserverBasics-WPF is a sample application that:

- Configures the Microsoft.Samples.Kinect.Webserver component to serve Kinect data on a localhost port
- Serves sample web page and associated static content
- Provides UI to start/stop the server and view errors encountered while serving data
- Writes informational messages to a log file

Redistributable web content (Kinect-1.8.0.js, KinectWorker-1.8.0.js and Kinect-1.8.0.css) provide:

- An API layer that abstracts communication with the server which web applications can use to get access to and manipulate Kinect data
- Button and cursor controls, with associated styles, which web applications can include in their overall user experience

#### Color Capture and Camera Pose Finder for Kinect Fusion

Kinect Fusion provides 3D object scanning and model creation using a Kinect for Windows sensor. With this release, users can scan a scene with the Kinect camera (now optionally also capturing low-resolution color) and simultaneously see, and interact with, a detailed 3D model of the scene. There is also support for recovering the camera pose once tracking has been lost, without having to find the very last tracked position or reset the whole reconstruction, by moving the camera close to one of the original camera positions during initial reconstruction. We also added an API to both the original depth reconstruction and new color reconstruction interfaces to match the "ExportVolumeBlock" and import a volume block. This is restricted to importing at the resolution of the created volume. Please see the feature documentation for an in depth discussion of the new functionality.

Kinect Color Fusion Samples:

- Kinect Fusion Color Basics-WPF, Kinect Fusion Color Basics-D2D: Demonstrate basic use of the Kinect Fusion APIs for 3D reconstruction with the option of using color.
- Kinect Fusion Explorer-WPF, Kinect Fusion Explorer-D2D: Similar to previous Kinect Fusion Explorer in 1.7 with an extra feature to additionally capture color while performing 3D reconstruction.

- Kinect Fusion Explorer Multi Static Cameras-WPF: This demonstrates having multiple static Kinect cameras integrate into the same reconstruction volume, given user-defined transformations for each camera.
- Kinect Fusion Head Scanning-WPF: Demonstrates how to leverage a combination of Kinect Fusion and Face Tracking to scan high resolution models of faces and heads.

#### Introducing New Samples!

- Adaptive UI-WPF: This new sample demonstrates the basics of adaptive UI that is displayed on screen in the appropriate location and size based on the user's height, distance from the screen, and field of view. The sample provides settings for interaction zone boundaries, tracks users and transitions as they move between far range and tactile range.
- Webservice Basics-WPF: This sample demonstrates how to use the Microsoft.Samples.Kinect.Webservice component to serve Kinect data on a localhost port. The component and sample require .NET 4.5 (Visual Studio 2012) and Windows 8 (or later).
- Background Removal Basics-D2D, Background Removal Basics-WPF: Demonstrates how to use the KinectBackgroundRemoval API. This is an improved version of the Coordinate Mapping sample (previously named Green Screen sample in 1.7).
- Kinect Fusion Explorer Multi Static Cameras-WPF: This demonstrates having multiple static Kinect cameras integrate into the same reconstruction volume, given user-defined transformations for each camera. A new 3rd person view and basic WPF graphics are also enabled to provide way for users to visually explore a reconstruction scene during setup and capture.
- Kinect Fusion Color Basics-D2D, Kinect Fusion Color Basics-WPF: Demonstrates the basics of Kinect Fusion for 3D reconstruction, now including low-resolution color capture.
- Kinect Fusion Head Scanning-WPF: Demonstrates how to leverage a combination of Kinect Fusion and Face Tracking to scan high resolution models of faces and heads.

#### Updated Samples

- Kinect Fusion Explorer-D2D, Kinect Fusion Explorer-WPF: Demonstrates additional features of Kinect Fusion for 3D reconstruction, now including low-resolution color capture. Please review the documentation for the hardware and software requirements. (Note: Suitable DirectX11 graphics card required for real-time reconstruction). The Explorer samples have been updated to always create a volume with the option of capturing color, hence GPU memory requirements have doubled compared to the v1.7 Explorer samples for the same voxel resolutions. Kinect Fusion Explorer-D2D also now integrates the Camera Pose Finder for increased robustness to failed tracking.
- Coordinate Mapping Basics-WPF, Coordinate Mapping Basics-D2D: These samples were previously named Green Screen in 1.7, renamed to Coordinate Mapping Basics in 1.8.

#### What's New in version 1.7 of the SDK and the Developer Toolkit

Here's a link to the [known issues](#) for this release.

This release provides many new features, all part of the Developer Toolkit 1.7.0. The SDK/Runtime v1.7 will contain minor changes.

#### Introducing new Kinect Interactions

We've built a new Interactions framework which provides *pre-packaged, reusable* components that allow for even more exciting interaction possibilities. These components are supplied in both native



and managed packages for maximum flexibility, and are also provided as a set of WPF controls. Among the new features are:

- **Press for Selection.** This provides, along with the new KinectInteraction Controls, improved selection capability and faster interactions. If you're familiar with previous Kinect for Windows interaction capabilities, this replaces the hover select concept.
- **Grip and Move for Scrolling.** This provides, along with the new KinectInteraction Controls, 1-to-1 manipulation for more precise scrolling, as well as large fast scrolls with a fling motion. If you're familiar with previous Kinect for Windows interaction capabilities, this replaces the hover scroll model.

New interactions work best with the following setup:

- User stands 1.5 - 2.0 meters away from the sensor
- Sensor mounted directly above or below the screen showing the application, and centered
- Screen size < 46 inches
- Avoid extreme tilt angles
- As always, avoid lots of natural light and reflective materials for more reliable tracking

### Engagement Model Enhancements

The Engagement model determines which user is currently interacting with the Kinect-enabled application.

This has been greatly enhanced to provide more natural interaction when a user starts interacting with the application, and particularly when the sensor detects multiple people. Developers can also now override the supplied engagement model as desired.

### APIs, Samples, and DLL Details

A set of WPF interactive controls are provided to make it easy to incorporate these interactions into your applications.

Two samples use these controls: ControlsBasics-WPF and InteractionGallery-WPF. The controls can also be installed in source form via Toolkit Browser -> "Components" -> Microsoft.Kinect.Toolkit.Controls.

- InteractionGallery - WPF uses the new KinectInteraction Controls in a customized app experience that demonstrates examples of navigation, engagement, article reading, picture viewing, video playback, and panning with grip. It was designed for 1920x1080 resolution screens in landscape layout.

For those building applications with UI technologies other than WPF, the lower level InteractionStream APIs (native or managed) are available to build on top of. Native DLLs for InteractionStream are `Kinect_Interaction170_32.dll` and `Kinect_Interaction170_64.dll` under `%KINECT_TOOLKIT_DIR%\Redist`. Managed DLL for InteractionStream is `Microsoft.Kinect.Toolkit.Interaction.dll` found in `%KINECT_TOOLKIT_DIR%\Assemblies`.

- There is no sample of InteractionStream API usage, however, `Microsoft.Kinect.Toolkit.Controls` source code (see info about controls samples above) is available and is a great example of using InteractionStream.

### Kinect Fusion

KinectFusion provides 3D object scanning and model creation using a Kinect for Windows sensor. The user can paint a scene with the Kinect camera and simultaneously see, and interact with, a detailed 3D model of the scene. Kinect Fusion can be run at interactive rates on supported GPUs, and can run at non-

interactive rates on a variety of hardware. Running at non-interactive rates may allow larger volume reconstructions.

### **Kinect Fusion Samples:**

- Kinect Fusion Basics - WPF, Kinect Fusion Basics - D2D: Demonstrates basic use of the Kinect Fusion APIs for 3D reconstruction.
- Kinect Fusion Explorer - WPF, Kinect Fusion Explorer - D2D: Demonstrates advanced 3D reconstruction features of Kinect Fusion, allowing adjustment of many reconstruction parameters, and export of reconstructed meshes.

### **Kinect Fusion Tech Specs**

Kinect Fusion can process data either on a DirectX 11 compatible GPU with C++ AMP, or on the CPU, by setting the reconstruction processor type during reconstruction volume creation. The CPU processor is best suited to offline processing as only modern DirectX 11 GPUs will enable real-time and interactive frame rates during reconstruction.

### **Minimum Hardware Requirements for GPU based reconstruction**

DirectX 11 compatible graphics card.

Kinect Fusion has been tested on the NVidia GeForce GTX560, and the AMD Radeon 6950. These cards, or higher end cards from the same product lines are expected to be able to run at interactive rates. For those building applications with technologies other than WPF, the lower level Fusion APIs (native or managed) are available to build on top of. Native DLLs for Fusion are Kinect\_Fusion170\_32.dll and Kinect\_Fusion170\_64.dll under *%KINECT\_TOOLKIT\_DIR%\Redist*. Managed DLL for Fusion is Microsoft.Kinect.Toolkit.Fusion.dll found in *%KINECT\_TOOLKIT\_DIR%\Assemblies*.

### **Recommended Hardware**

Desktop PC with 3GHz (or better) multi-core processor and a graphics card with 2GB or more of dedicated on-board memory. Kinect Fusion has been tested for high-end scenarios on a NVidia GeForce GTX680 and AMD Radeon HD 7850.

Note: It is possible to use Kinect Fusion on laptop class GPU hardware, but this typically runs significantly slower than desktop-class hardware. In general, aim to process at the same frame rate as the Kinect sensor (30fps) to enable the most robust camera pose tracking.

### **Getting Started with Kinect Fusion (Important!)**

1. Ensure you have compatible hardware (see Tech Specs section above).
2. Download and install the latest graphics display drivers for your GPU.

### **Kinect Sensor Chooser - native**

The Kinect Sensor Chooser is a native component that allows simplified management of the Kinect Sensor lifetime, and an enhanced user experience when dealing with missing sensors, unpowered sensors, or sensors that get unplugged while an application is running. It provides similar capabilities to the KinectSensorChooser in the Microsoft.Kinect.Toolkit component. A NuiSensorChooserUI control is also provided for use in native applications. It provides a user experience similar to the managed KinectSensorChooserUI.

### **Introducing New Samples!**

- Controls Basics - WPF: Demonstrates the new KinectInteraction Controls, including hands-free button pressing and scrolling through large lists. This replaces the Basic Interactions sample from previous releases.

- Interaction Gallery - WPF: Demonstrates basic interactions using the new KinectInteraction Controls.
- KinectBridge with MATLAB Basics - D2D: Demonstrates how to do image processing with the Kinect sensor using MATLAB API.
- KinectBridge with OpenCV Basics - D2D: Demonstrates how to do image processing with the Kinect sensor using OpenCV API.
- Kinect Explorer - D2D: Demonstrates how to use the Kinect's ColorImageStream, DepthImageStream, SkeletonStream, and AudioSource with C++ and Direct2D. This replaces the SkeletalViewer C++ sample.
- Kinect Fusion Basics - WPF, Kinect Fusion Basics - D2D: Demonstrates basic use of the Kinect Fusion APIs for 3D reconstruction.
- Kinect Fusion Explorer - WPF, Kinect Fusion Explorer - D2D: Demonstrates advanced 3D reconstruction features of Kinect Fusion, allowing adjustment of many reconstruction parameters, and export of reconstructed meshes.

What's New in version 1.6 of the SDK and the Developer Toolkit

Here's a link to the [known issues](#) for this release.

#### Windows 8 Support

Using the Kinect for Windows SDK, you can develop a Kinect for Windows application for a desktop application in Windows 8.

#### Visual Studio 2012 Support

The SDK supports development with Visual Studio 2012, including the new .NET Framework 4.5.

#### Accelerometer Data APIs

Data from the sensor's accelerometer is now exposed in the API. This enables detection of the sensor's orientation.

#### Extended Depth Data Is Now Available

CopyDepthImagePixelDataTo() now provides details beyond 4 meters; please note that the quality of data degrades with distance. In addition to Extended Depth Data, usability of the Depth Data API has been improved. (No more bit masking is required.)

#### Color Camera Setting APIs

The Color Camera Settings can now be optimized to your environment.

- You can now fine-tune white balance, contrast, hue, saturation, and other settings.
- To see the full list, launch Kinect Explorer from Developer Toolkit Browser and review the **Exposure Settings** and **Color Settings** controls for a full list of settings that can be optimized.

#### More Control over Decoding

New RawBayer Resolutions for ColorImageFormat give you the ability to do your own Bayer to RGB conversions on CPU or GPU.

#### New Coordinate Space Conversion APIs

There are several new APIs to convert data between coordinate spaces: color, depth, and skeleton. There are two sets of APIs: one for converting individual pixels and the other for converting an entire image frame.

### German Language Pack for Speech Recognition

The SDK ships with a German speech recognition language pack that has been optimized for the sensor's microphone array.

### Infrared Emitter Control API

The sensor's infrared emitter has previously always been on when the sensor is active, which can cause depth detection degradation in a scenario where multiple sensors are observing the same space. There is a new API (`KinectSensor.ForceInfraredEmitterOff`) for turning the infrared emitter off.

### Introducing New Samples!

- **Basic Interactions-WPF:** Demonstrates basic gestures, such as targeting and selecting with a cursor, as well as appropriate feedback mechanisms for an optimal user experience.
- **WPF D3D Interop:** Demonstrates DirectX 11 interop with WPF, including full WPF composition of DirectX surfaces.
- **Infrared Basics-WPF, Infrared Basics-D2D:** Demonstrates using an infrared stream and displaying an image using depth data.

### Kinect Studio 1.6.0

Kinect Studio has been updated to support the Infrared, RawBayer, Extended Depth Data, and Accelerometer features.

### The Infrared Stream Is Now Exposed in the API

The Kinect sensor's infrared stream is now exposed as a new `ColorImageFormat`. You can use the infrared stream in many scenarios, such as:

- Calibrating other color cameras to the Kinect's depth sensor
- Capturing grayscale images in low-light situations

Two infrared samples have been added to the toolkit, and you can also try out infrared in `KinectExplorer`.

The sensor is not capable of capturing an infrared stream and a color stream simultaneously. However, you can capture an infrared stream and a depth stream simultaneously.

### Support for Virtual Machines

The Kinect for Windows sensor now works on Windows running in a virtual machine and has been tested with the following VM environments:

- Microsoft Hyper-V
- VMWare
- Parallels

This greatly expands the utility of the Kinect for Windows SDK, as it can now be used on any machine whose native OS supports running Windows in one of the VM packages listed above. In particular, this enables several developer scenarios, such as certain automated testing approaches.

Setup and configuration details for using the Kinect with the tested VMs are contained in the `Getting Started` section of this documentation.

Note that only one Kinect at a time will work with a given VM, and you may experience lower frame rates on lower-end computers as some computing resources are consumed by the VM itself.

---

### What's New in version 1.5.2 of the Developer Toolkit

Here's a link to the [known issues](#) for this release.

### New Interoperability Sample

WPF3DInterop, a new sample that demonstrates DirectX 11 interoperability with Windows Presentation Foundation (WPF), including full WPF composition of DirectX surfaces. This enables powerful DirectX rendering intermixed with quicker-to-develop WPF user interfaces.

---

### What's New in version 1.5.1 of the Developer Toolkit

Here's a link to the [known issues](#) for this release.

#### Improvements to Kinect Studio 1.5.1

We made the following performance and stability improvements to Kinect Studio 1.5.1:

- Reduced CPU usage overhead.
- Improved toolbar icons in Normal and High Contrast mode.
- Eliminated distortion and flickering that occurred in the Color Viewer display during playback using some video cards.
- Fixed hanging that occurred using AMD Radeon HD 5xxx series graphics cards.
- Fixed crashing that occurred during playback of .xed files that were recently saved to a network share.

#### Improvements to Face Tracking

We made the following improvements to the Face Tracking feature:

- The Microsoft.Kinect.Toolkit.FaceTracking project now automatically copies the appropriate native libraries to the output directory in a post-build step.
- The FaceTracking3D-WPF sample now builds all configurations.

#### Sample Updates

We made the following fixes to the samples:

- Audio Explorer – Fixed label captions that were clipped.
- Avateering – Reference to Microsoft.Kinect.dll is no longer version-specific.
- Kinect Explorer – Fixed crash that occurred in locales where "," is used as the numeric separator.

#### Offline Documentation

SDK documentation is now available offline. For details see [Kinect for Windows SDK Offline Docs](#).

#### Windows 8

We tested SDK 1.5 and Toolkit 1.5.1 on top of Windows 8 Release Preview and found no new problems. Although that is not yet a supported production platform for Kinect for Windows, we continue to track well towards supporting it sometime after Windows 8 ships.

---

### What's New in version 1.5 of the SDK and the Developer Toolkit

Here's a link to the [known issues](#) for this release.

#### Backward Compatibility

The Kinect for Windows 1.5.0 SDK, driver, and runtime are 100% compatible with Kinect for Windows 1.0 application binaries.

#### Kinect for Windows Developer Toolkit

Please note the following changes to the developer toolkit:

- As of this release, the SDK has been divided into a core SDK and a developer toolkit. These are two separate installations.
- All samples have been moved into the toolkit.
- We've continued significant sample investments in SDK 1.5.0. There are many new samples in both C++ and C#. In addition, we've included a "Basics" series of samples with language coverage in C++, C#, and Visual Basic. To explore the list of new and updated samples, please launch the Developer Toolkit Browser and explore.
- We've taken KinectSensorChooser, formerly part of the WpfViewers, and split the logic and UI into two different classes: KinectSensorChooser and KinectSensorChooserUI in Microsoft.Kinect.Toolkit.dll.
- KinectSensorChooser can be used in non-WPF scenarios because it provides logic only and has no UI.
- We have significantly improved the user experience for KinectSensorChooserUI (used with a KinectSensorChooser instance). The ShapeGame sample now uses it.

#### Kinect Studio

Kinect Studio is a new tool that allows you to record and replay Kinect data to aid in development. For example, a developer writing a Kinect for Windows application for use in a shopping center can record clips of users in the target environment and then replay those clips at a later time for development and test purposes.

Please note the following when using Kinect Studio:

- Kinect Studio must be used in conjunction with a Kinect for Windows application. Start the application, then start Kinect Studio, and you can then record and replay clips. When you replay a clip, it will be fed into the application as if it were live Kinect data.
- Kinect Studio puts additional load on your machine and it may cause the frame rate to drop. Using a faster CPU and memory will improve performance.
- Your temporary file location (set using **Tools > Options**) should not be a network location.
- For more information, see [1.5.0 SDK and Developer Toolkit Known Issues](#).

#### Skeletal Tracking in Near Range Now Available

It is now possible to receive skeletal tracking data when the Kinect camera is in the Near Range setting. When using near-range skeletal tracking, please note the following:

- This setting is disabled by default to ensure backward compatibility with Kinect for Windows 1.0 applications.
- Enable this setting through the SkeletonStream.EnableTrackingInNearRange property (in managed code) or by including the NUI\_SKELETON\_TRACKING\_FLAG\_ENABLE\_IN\_NEAR\_RANGE flag when calling NuiSkeletonTrackingEnable (in native code).
- This feature works whether SkeletonTrackingMode is set to Default or Seated.
- We suggest using Seated mode when using Near Range, because in most scenarios the player body will not be entirely visible. Ensure the torso and the head of the players are visible for locking.

#### Seated Skeletal Tracking Now Available

Skeletal tracking is now available when Seated mode is selected.

- This mode tracks a 10-joint skeleton (consisting of head, shoulders, and arms) and ignores the leg and hip joints.
- This mode can be used regardless of whether the player is standing or sitting.
- Seated mode has been added to the Skeletal Viewer (C++) and Kinect Explorer (C#) samples. To try out the mode to understand its tracking ability, launch one of those applications and change the tracking setting from Default to Seated.
- For information on enabling Seated mode, see [Natural User Interface for Kinect for Windows](#).
- Seated mode skeletal tracking has higher performance requirements than Default mode, especially when tracking two players. You may notice a reduction in frame rate with some system configurations.

### Runtime Improvements

We have made the following improvements to the runtime:

- The performance of the `KinectSensor.MapDepthFrameToColorFrame` method has been significantly improved. It is now five times faster on average.
- Depth and color frames are now kept in sync. The Kinect for Windows runtime continuously monitors the depth and color streams and ensures that there is minimal drift between them.
- In managed code you will see that the frames returned from the `KinectSensor.AllFramesReady` event will have been captured at nearly the same time and will have timestamps that verify this.

### RGB Image quality

We have made the following improvements to image quality:

- The RGB camera stream quality has been improved for the RGB 640x480/30fps and YUV 640x480/15fps video modes.
- The image quality is now sharper and more color-accurate in high and low lighting conditions.

### Joint Orientation

Kinect for Windows runtime provides joint orientation information for the skeletons tracked by the ST pipeline. The joint orientation is provided in two forms:

- A hierarchical rotation based on a bone relationship defined on the ST joint structure
- An absolute orientation in Kinect camera coordinates.

The orientation information is provided in the form of quaternions and rotation matrices. This information can be used in avatar animation scenarios as well as simple pose detection.

### New Supported Languages for Speech Recognition

Acoustic models have been created to allow speech recognition in several additional locales. These are runtime components that are packaged individually and are available [here](#). The following locales are now supported:

- en-AU
- en-CA
- en-GB
- en-IE
- en-NZ
- es-ES
- es-MX
- fr-CA

- fr-FR
- it-IT
- ja-JP

#### Face Tracking SDK

We have added a Face Tracking component, which offers the following features:

- The Face Tracking component tracks face position, orientation, and facial features in real time.
- A 3D mesh of the tracked face, along with eyebrow position and mouth shape, is animated in real time.
- Multiple faces can be tracked simultaneously.
- Face Tracking components can be used in native C++. A managed wrapper is provided for C# and VB projects.

#### Documentation Improvements

We made the following improvements to the documentation:

- The documentation is now online in MSDN Library. To view the documentation, press F1 in Visual Studio.
- The SDK Documentation .chm file is no longer distributed by setup. Please use the [online documentation](#).



## **Anexos Digitales**