

ADiBAS: Sistema Multisensor de Adquisición Automática de Datos Corporales Objetivos, Robustos y Fiables para el Análisis de la Postura y el Movimiento

Miguel Reyes^{1,3}, José Ramírez-Moreno², Juan R. Revilla², Petia Radeva^{1,3},
y Sergio Escalera^{1,3}

¹ Depart. De Matemàtica Aplicada i Anàlisi, Universitat de Barcelona, Gran Via de les Corts Catalanes 585, 08007, Barcelona. e-mail: mreyese@gmail.com

² Instituto de Fisioterapia Global Mezières, Guillem Tell 40 Entlo. 2ª, 08006 Barcelona. e-mail: {joseramirez,jrevilla}@kinemez.com

³ Centro de Visión por Computador, Campus Universitat Autònoma de Barcelona, Edifici O, 08193, Bellaterra, Barcelona. e-mail: {petia,sergio}@maia.ub.es

Resumen.

El análisis de la postura y del rango de movimiento son fundamentales para conocer la optimización del gesto y mejorar, de este modo, el rendimiento y la detección de posibles lesiones. Esta cuantificación es especialmente interesante en deportistas o en pacientes que presentan alguna lesión neurológica o del sistema musculoesquelético, ya que permite conocer el proceso evolutivo de estos pacientes, evaluar la eficacia de la terapia aplicada y proponer, en caso necesario, una modificación del protocolo de tratamiento. En este trabajo presentamos un sistema automático que permite, mediante una tecnología no invasiva, la captación automática de marcadores LED situados sobre el paciente y su posterior análisis con el fin de mostrar al especialista datos objetivos que permitan un mejor soporte diagnóstico. También se describe un sistema analítico de la postura corporal sin marcadores, donde su ejecución durante secuencias dinámicas aporta un alto grado de naturalidad al paciente a la hora de realizar los ejercicios funcionales.

Palabras clave: Análisis Corporal, Visión por Computador, Biometría Digital, Análisis Automático Objetivo.

1. Introducción

Durante el proceso de la extracción de la información referente a la postura y el gesto debemos realizar una representación objetiva, precisa y máximamente descriptiva para posteriormente elaborar un análisis valorativo por parte del experto. Las prácticas actuales para analizar la postura de forma automática son caras e invasivas y en algunos casos requieren de voluminosas instalaciones.

El problema de obtener una información precisa y fiable de la postura corporal es debido a la morfología móvil, no rígida y articulada de la que está compuesto el cuerpo humano. De esta forma es difícil encontrar una herramienta de medición que se adapte a las diferentes y múltiples configuraciones que puede adoptar la postura corporal. Es en este momento donde aparece la biometría digital. La biometría digital se define por la Sociedad Americana de fotogrametría como: "la tecnología para obtener información fiable de los objetos físicos o del entorno, a través del registro de las imágenes, se realiza la medición o interpretación". Los sistemas basados en esta tecnología son capaces de estimar alteraciones morfológicas o funcionales, siendo un acceso recurso para los profesionales de los expertos de la salud. La solución mayormente adoptada es la utilización de una instalación con múltiples cámaras [1,2] aplicando el concepto de visión

estéreo sobre una previa configuración y calibración de las cámaras. Este sistema es costoso, voluminoso y no portable. Además, suele requerir de ciertas condiciones de iluminación muy restringidas, aunque tienen la ventaja de ser poco invasivos, ya que suelen utilizar pequeños marcadores ligeros reflectantes. La alternativa a este sistema comúnmente más utilizada es el uso de acelerómetros [3]. Esta solución también suele ser costosa e invasiva, ya que éstos suelen ser de grandes dimensiones e incluso no pueden colocarse en cualquier parte del cuerpo ni obtener mediciones espaciales en posiciones estáticas. Además, suelen dar resultados poco precisos debido a la complejidad de las articulaciones multiaxiales.

1.1 Aportación

En este trabajo presentamos ADiBAS, un sistema de medición robusto sobre las diferentes localidades del cuerpo humano, siendo económico y fácilmente personalizable a las necesidades de análisis de los diferentes usuarios. Una de las principales ventajas del sistema propuesto es ser mínimamente invasivo.

En el campo de la actividad física y el deporte, tanto para el profesional como para el amateur, es importante conocer la dinámica del gesto para optimizar los recursos energéticos, los resultados y prevenir futuras lesiones tan frecuentes entre los deportistas. La magnitud de las velocidades y aceleraciones angulares asociadas a las actividades deportivas son captadas por el sistema propuesto a través de la identificación automática de los LEDs situados en el cuerpo del deportista y/o en el aparato utilizado. Así por ejemplo, deportes como el ciclismo en el que se puede pedalear durante muchas horas, es importante la adecuada adaptación de la posición del ciclista en la bicicleta si se quieren evitar posibles lesiones. En otras actividades como andar, correr, saltar, lanzar, chutar, etc., la captura del movimiento proporciona una herramienta muy útil para el análisis clínico o con fines de investigación. De este modo, se pueden conocer los efectos de ciertos aparatos ortopédicos, instrumentos propios de cada deporte, la mejora en el rendimiento, la prevención de lesiones o la reincorporación a la actividad del deportista.

En rehabilitación es importante mantener la atención y participación del paciente. Esto se consigue gracias a la interfaz desarrollada para cada patología. La visualización en la pantalla del ordenador del movimiento en tres dimensiones proporciona un entorno de retroalimentación específico para cada paciente. Este biofeedback utilizado por nuestro software es especialmente relevante porque, de este modo, se produce una mayor integración sensitivo-motriz, una mejora de la neuroplasticidad, una mayor eficacia en el rendimiento del gesto y, en consecuencia, una recuperación más precoz, más corta y más amena. De la misma forma, el sistema de análisis en tiempo real que proporciona nuestro software ofrece al especialista una información útil sobre pacientes con problemas neurológicos tales como: accidentes cerebrovasculares, traumatismos craneales, enfermedades con alteración del equilibrio; pacientes geriátricos para la prevención de caídas, ejercicios de activación motora; traumatología y reumatología en deformidades de la columna vertebral, intervenciones quirúrgicas protésicas, lesiones músculo-esqueléticas, etc.

2. Metodología

El sistema propuesto se compone de un multisensor y un ordenador de sobremesa o portátil. El multisensor contiene un láser infrarrojo y una cámara de vídeo RGB. La cámara de vídeo usa SVGA de 8 bits a 30Hz, con una resolución de imagen de 640x480píxeles. El láser de infrarrojos se combina con un sensor CMOS



Figura 1. Elementos principales del sistema

monocromo que captura datos 3D del entorno. Para obtener los datos 3D, el sensor infrarrojo dispara una nube de puntos contra el entorno. Cada píxel de profundidad es calculado entonces a través de la diferenciación de los píxeles distribuidos a más resolución. Este valor es inversamente proporcional al radio de la Gaussiana sobre cada punto, el cual es linealmente proporcional a la profundidad real de cada píxel. La sensibilidad del sensor ha sido testada en un rango de entre 1 a 3,5 metros, con un angular de de 57 y 43 grados en los ejes horizontal y vertical, respectivamente. En relación a la implementación del sistema que gestiona la comunicación con los sensores, se han usado las librerías de libre distribución OpenNI y OpenCV, dentro de un entorno gráfico de programación de código libre Qt basado en lenguaje C++. El testeo del sistema se ha realizado en PCs ordinarios de capacidades estándar: sistema operativo Windows Vista/Windows 7 de 32 bits, con un equipamiento hardware de procesador Intel i3 y 2GB de memoria RAM. El dispositivo multisensor utilizado se muestra a la derecha de la Figura 1.

Una de las ventajas de utilizar un sistema multimodal dotado de una cámara RGB y sensores de profundidad (D, del inglés *Depth*) es la obtención del espacio RGBD a través de la alineación de ambas, como se puede observar en la Figura 2. El análisis de este espacio nos permitirá realizar mediciones tridimensionales fiables a través de técnicas de visión por computador y técnicas de fotogrametría. Una de las ventajas del sistema adoptado es la opción de realizar cualquier medición del objetivo de las cámaras alineadas, ya que sobre cada punto de resolución que observa el usuario se poseen las coordenadas mundo en milímetros.



Figura 2. Visualización del espacio RGBD

2.2 Fases del sistema

El desarrollo del proyecto se puede dividir principalmente en dos etapas. Una primera etapa basada en técnicas de fotogrametría digital, las cuales nos permitirán realizar medidas reales sobre imágenes enfocadas en ángulo libre, es decir, realizaremos medidas tridimensionales precisas sin tener que colocar la cámara de una forma específica, tan solo enfocando el objetivo a realizar las medidas. Esta libre localización de la posición de la cámara nos permite una gran portabilidad del sistema, permitiendo su utilización en escenarios de diferentes dimensiones. Aunque el usuario no debe realizar ninguna calibración para obtener medidas precisas, el proceso de calibración es realizado previamente con el fin de obtener los parámetros intrínsecos de la cámara y los diferentes coeficientes de distorsión de la cámara. La segunda fase del desarrollo del proyecto se refiere al proceso de automatización de captación de marcadores, donde intervienen técnicas de visión por computador. Esta fase de automatización dota de novedad al sistema, ya que ninguno de los sistemas de biometría digital examinados detecta los marcadores automáticamente, sino que el usuario debe inicializar la posición de cada marcador. Creemos que la detección automática es una de las características principales del sistema, proporcionando un ahorro de tiempo al usuario del sistema, como la posibilidad de establecer y asociar protocolos de postura personalizados por el experto en salud.

Calibración: Pretendiendo realizar mediciones con un alto grado de precisión primeramente debemos corregir las distorsiones de la cámara. Existen tres tipos de distorsiones en fotogrametría: la distorsión esferoide, un efecto de lente en el que las líneas rectas se curvan hacia los bordes, la distorsión en corsé, teniendo el efecto opuesto, y distorsión por descentramiento debido a la falta de coincidencia de los centros de curvatura de las lentes con el eje óptico del objetivo, es decir, a la falta de colinearidad de los centros.

Tanto para obtener estos coeficientes como para los parámetros intrínsecos de la cámara, los cuales modelan la geometría interna de la cámara y las características ópticas de la lente, utilizamos una plantilla de calibración cuadriculada de medidas preestablecidas. A través de restricciones epipolares y trilineales entre pares y tripletas de imágenes se estiman los coeficientes de distorsión [4], en cuanto para la obtención de los puntos de la plantilla se utiliza un algoritmo de extracción esquinas muy preciso [5]. Una vez corregida la lente, son necesarios los parámetros intrínsecos de la cámara, los cuales determinarán las coordenadas de un punto en la imagen a partir de la posición del punto en la escena respecto del sistema de coordenadas de la cámara, es decir, el modelo pin-hole [6].

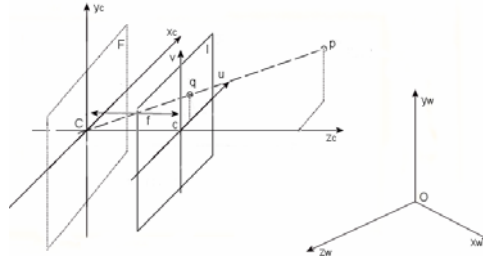


Figura 3. Representación geométrica del modelo PIN-HOLE

Para el modelo pin-hole de la cámara se va a tomar el punto c como origen de medidas en la imagen. Se fija un sistema de coordenadas 3D para el escenario ($O, \{x_w, y_w, z_w\}$), otro sistema de coordenadas 3D para la cámara ($O', \{x_c, y_c, z_c\}$) y otro 2D en el plano de la imagen ($c, \{u, v\}$). El eje óptico está alineado con el eje Z del sistema de referencia de la cámara tal como se observa en la Figura 3. Las coordenadas de los puntos de interés en el espacio están referidas al sistema de coordenadas del escenario ($O, \{x_w, y_w, z_w\}$) y sus correspondientes posiciones en la imagen están referidas respecto al sistema de coordenadas de la misma ($c, \{u, v\}$). Para conseguir las coordenadas mundo, la solución se base en hallar la transformación lineal entre las coordenadas cámara a través de los coeficientes de los valores intrínsecos y las coordenadas mundo [7]. Una vez obtenidos los valores intrínsecos, para obtener una coordenada mundo utilizamos la ecuación 1.

$$x_w = \frac{(u - x_c) * z_w}{f_x}; y_w = \frac{(v - y_c) * z_w}{f_y} \quad (1)$$

Como último apartado referente a la fotogrametría del sistema, cabe destacar la alineación entre la cámara RGB y la cámara de profundidad. Debido a que en nuestro hardware multisensor vienen incorporadas juntas, y aunque el middleware OpenNI realiza una alineación bastante precisa, existen zonas donde la alineación no es del todo precisa para cubrir nuestros objetivos de medición. Para ello necesitamos una matriz de transformación capaz de interpolar la posición de punto de la cámara RGB sobre el sensor de profundidad. Con este fin, construimos una matriz de homografía H a través del método no lineal de mínimos cuadrados [8]. Para la composición de esta transición se realiza una recolección manual de 100 puntos correspondientes entre ambas cámaras.

Detección de marcadores. Una de las principales aportaciones del sistema desarrollado es la detección automática de marcadores corporales. Cada marcador está compuesto por una pila botón recargable IR2032 de 20 mm de diámetro con su correspondiente portapila, un microinterruptor y un LED de 3mm color blanco frío, 6000-8000mcd, 3V y 20mW. Este ensamblado permite una duración de alta luminosidad máxima de 2 horas. Con el fin de evitar tener que acondicionar la iluminación del escenario, se incorpora al sistema multisensor, externamente, un filtro infrarrojo digital de 850nm. Esta configuración permite captar los marcadores LED de forma instantánea, debido a que la luz infrarroja es altamente diferenciable del resto del escenario. Estos elementos del sistema se muestran a la izquierda de la Figura 1. Para evitar posibles falsos positivos del escenario, se pre-procesa la imagen, ejecutando primitivas de suavizado y eliminación de

valores tenues y cambio de escala de color a HSL. Se utiliza un espacio de color diferente al RGB, en este caso HSL, por su mayor discriminación entre colores, aumentando la precisión de detección y disminuyendo ambigüedades. Posteriormente, se analizan las formas y siluetas resultantes, descartando aquellas que sean incoherentes con la silueta de un marcador LED, obteniendo su centro de masa. Este centro es el que mantendremos como referencia del marcador LED para el análisis global. Una vez recogidos todos los marcadores corporales, el sistema compara la estructura de puntos extraída de las configuraciones personalizadas y guardadas previamente por el experto, a las que nombramos protocolos de postura. Cuando el sistema encuentre coherencia de estructura postural con un protocolo, éste sólo mostrará aquellas medidas y ángulos de interés requeridos por el protocolo.

3. Resultados y discusión

Para la validación del sistema propuesto analizamos dos características: la captación de marcadores y la precisión de las mediciones tridimensionales (tanto en distancias milimétricas como en grados para cada tripleta de puntos). Cabe destacar que se ha realizado un protocolo de validación exhaustivo, mediante la utilización de herramientas de medición certificadas y homologadas, así como la participación de diferentes examinadores. Para el estudio de validación el sistema se ha diseñado un conjunto de pruebas altamente variado en configuraciones de posición. Para reflejar una situación lo más realista posible, el escenario de validación es un centro asistencial de rehabilitación funcional. Los marcadores corporales se han colocado simulando un paciente de dimensiones estándar.

Principalmente el sistema multisensor se evalúa en posición horizontal y vertical. El hecho de probar el sistema en posición vertical es debido a que obtiene una mayor resolución horizontal, propiedad aprovechable en cuanto al análisis de sujetos en posición vertical. Una de las características del sistema que estamos interesados en analizar es su rendimiento bajo diferentes distancias entre sujeto y dispositivo multisensor. Bajo las diferentes distancias, se abre un abanico de combinaciones a partir de la rotación sobre los ejes verticales y horizontales del sistema. Se ha diseñado un protocolo de captación y medición basado en la colocación de cuatro triángulos. El objetivo de cada prueba es localizar cada uno de los tres vértices, para posteriormente obtener las distancias y ángulos entre los vértices de cada triángulo. Las combinaciones realizadas sobre cada una de las pruebas consisten en rotaciones de 20 grados en sentido este y oeste sobre el plano el eje horizontal del dispositivo multisensor, rotaciones de 20 grados en sentido norte y sur sobre el plano vertical, y desplazamientos de 50 cm del dispositivo sobre el plano perpendicular al objetivo.

DISTANCIAS LONGITUDINALES						
Distancia Kinect – Objeto	1,3 metros		1,9 metros		2,2 metros	
Posición Kinect	H	V	H	V	H	V
Captación Marcadores	6/6	9/9	9/9	12/12	12/12	12/12
Promedio error absoluto mm	1.07	1.55	1.55	2.48	2.99	3.18
Promedio error relativo %	0.54	0.84	0.84	1.41	1.75	1.94
Standard Error Measurement %	1.12	1.44	1.44	1.01	2.22	1.01

Tabla 1. Resultados de la fase de validación referentes a longitudes

ANGULOS						
Distancia Kinect – Objeto	1,3 metros		1,9 metros		2,2 metros	
Posición Kinect	H	V	H	V	H	V
Captación Marcadores	6/6	9/9	9/9	12/12	12/12	12/12
Promedio error absoluto°	0.51	0.75	0.75	1.04	1.24	1.15
Promedio error relativo %	0.99	1.38	1.38	2.06	2.03	1.99
Standard Error Measurement %	0.36	0.47	0.47	0.90	0.77	0.90

Tabla 2. Resultados de la fase de validación referentes a ángulos

Uno de los primeros rasgos que se observan en las Tablas 1 y 2 de los resultados obtenidos, es que la precisión del sistema varía dependiendo de la distancia entre el sujeto y el dispositivo multisensor. Observamos que a mayor proximidad, mejor precisión obtenemos en las medidas tridimensionales. Otra característica referente a la detección de marcadores, es que necesitamos una distancia de 1.9 metros para poder detectar la totalidad de estos, que es el caso más semejante a una persona real. En cuanto a la propia detección, aseguramos que el método es suficientemente robusto ya que en todas las pruebas se han detectado todos los marcadores que estaban dentro del campo de visión del sistema multisensor. Finalmente en la Figura 4 se muestran ejemplos de detección automática de marcadores para diferentes protocolos de postura y número de puntos. Se puede apreciar la robustez en la detección. En la interfaz del sistema se muestra la distribución numerada de los puntos detectados, y todas las métricas obtenidas (en distancias y ángulos) se formatean y guardan en hojas de cálculo para su posterior análisis por parte del especialista. Con la intención de analizar en profundidad los datos obtenidos, comparamos nuestros resultados con trabajos recientes del mismo tópico [9], en el cual utilizan un sistema multisensor de altas prestaciones y con una inicialización manual de los marcadores. También se compara nuestro sistema con los sistemas multicámara [1] anteriormente comentados. Los valores de la comparativa se muestran en la Tabla 3. En [1] se describe uno de los trabajos más avanzados y recientes del estado del arte sobre sistemas multicámara. La metodología empleada de [1] es utilizada en grandes producciones cinematográficas dada su elevada precisión. Esta alta precisión se consigue mediante la combinación de resultados de múltiples cámaras. La comparativa se realiza en función de la precisión en obtener ángulos tridimensionales sobre estructuras articuladas.

Método	Error medio de métrica
Measurement of Angles in Time-of-Flight Data for the Automatic Supervision of Training Exercises	15°
The Multiple-Camera 3-D Production Studio 3 camaras	2,3 cm
The Multiple-Camera 3-D Production Studio 9 camaras	1,2 cm
The Multiple-Camera 3-D Production Studio 27 camaras	1.5 mm
The Multiple-Camera 3-D Production Studio 45 camaras	<0.5mm

Tabla 3. Comparativa de métodos de análisis postural

El caso de más fácil comparativa es con el trabajo presentado en [9], ya que se ha testeado sobre un contexto muy parecido al nuestro. En cambio, el trabajo [1], no ha sido validado sobre un contexto de realización de mediciones, sino el objetivo del sistema es realizar un modelado tridimensional de los objetos del escenario lo más realista posible, usando una metodología mucho más invasiva y costosa. En la figura 4 podemos observar algunos resultados obtenidos sobre casos reales. En la figura 4 a) y b) se analiza la postura de un paciente real sobre una vista frontal, mientras que en d) y e) se analiza el plano perfil del paciente. La figura 4 c) forma parte del conjunto de pruebas para completar la fase de validación.

4. Conclusiones y trabajo futuro

En este trabajo hemos presentado ADiBAS, un sistema de medición tridimensional capaz de obtener información métrica fiable de la postura corporal. Analizando los valores obtenidos de la fase de validación demostramos que estamos ante un sistema robusto, capaz de detectar los marcadores corporales bajo diferentes configuraciones de instalación del escenario. Destacamos que los valores obtenidos de las métricas de la fase de validación, en términos de detección automática de marcadores, distancias y ángulos, son lo suficientemente fiables como para elaborar estudios clínicos sobre pacientes reales. También cabe destacar que el sistema final obtenido cumple tanto con los requisitos económicos como de portabilidad, siendo no invasivo, lo que hace del sistema final un elemento práctico y de fácil instalación para trabajar en diferentes campos del análisis de la salud física.

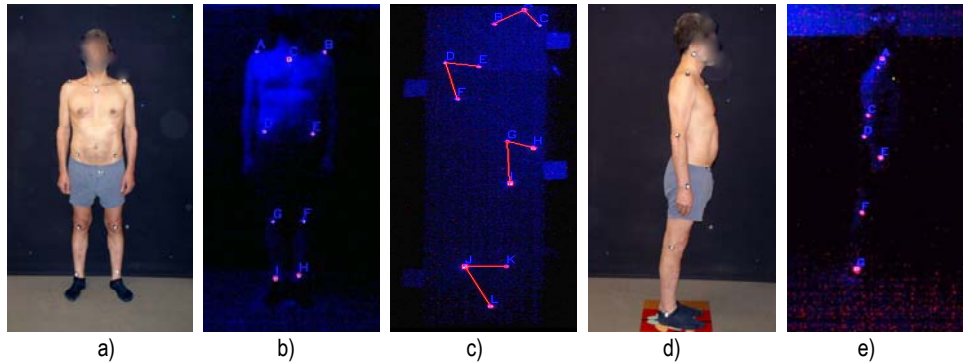


Figura 4 a), b), d), y e) Casos prácticos de la detección de marcadores sobre sujetos reales, c) Estructura de los marcadores en el protocolo de validación

Durante el desarrollo del sistema se han detectado ciertas restricciones que se deben cumplir para obtener fiabilidad en los datos. Ya que nuestro sistema es similar en metodología a los sistemas basados en multicámara, es susceptible a las mismas situaciones críticas, como por ejemplo las oclusiones. Se debe tener en cuenta que aquellas zonas que quedan ocluidas al objetivo del dispositivo multisensor no dispondrán de datos de medición fiables. Por lo tanto, una restricción para elaborar una medición fiable es que todos los marcadores corporales sean visibles por el dispositivo de captura.

Dado que la detección de marcadores corporales es efectiva, el software adquiere un mayor acercamiento al experto, mostrando los datos de interés previamente configurados por éste. Este aporte dota al sistema de síntesis analítica y ahorro de tiempo al experto en salud. Esta característica provoca que el software sea altamente adaptable a las necesidades de cada paciente, abriendo así un amplio abanico de aplicaciones del sistema.

Una de las líneas actuales en desarrollo consiste en el análisis automático inteligente de las trayectorias de movimiento de los marcadores de forma que se puedan hacer análisis automáticos sobre el rango de movimiento de los sujetos.

Referencias

- [1] Jonathan Starck, Atsuto Maki, Shohei Nobuhara, Adrian Hilton, and Takashi Matsuyama, The Multiple-Camera 3-D Production Studio, IEEE Transactions on circuits and systems for video technology, vol 19, no. 6, June 2009.
- [2] Ivana Mikić, Mohan Trivedi, Edward Hunter, Pamela Cosman, Articulated Body Posture Estimation from Multi-Camera Voxel Data, Computer Vision and Robotics Research Lab Department of Electrical and Computer Engineering University of California, San Diego, La Jolla, CA 92093.
- [3] Kourosh Parsa, Jorge Angeles, and Arun K. Misra, Department of Mechanical Engineering and Centre for Intelligent Machines. McGill University, Montreal. Canada. Proceedings of the 2001 IEEE. International Conference on Robotics & Automation. Seoul, Korea. May 21-26, 2001.
- [4] Stein G. (1997) Lens distortion calibration using point correspondences. Proceedings of the International Conference on Computer Vision and Pattern Recognition p 602-608.
- [5] Konstantinos G. Derpanis. Harris Corner Detector. International Journal of Computer Vision, October 27, 2004.
- [6] Andrea Busello, PIN-HOLE Camera Geometry, <http://homepages.inf.ed.ac.uk/cgi/rbf/CVONLINE/>, June 2002.
- [7] Zhang Z. (1998) A flexible new technique for camera calibration. Technical Report MSR-TR-98-71, Microsoft Research
- [8] Homography Estimation, David Kriegman, Computer Vision I CSE 252A, Winter 2007.
- [9] S. Soutschek, J. Kornhuber Psychiatrische und Psychotherapeutische Klinik, Universitätatsklinikum Erlangen, Germany A. Maier, S. Bauer, P. Kugler, J. Hornegger Lehrstuhl für Mustererkennung, Universität at Erlangen-Nürnberg, Germany M. Bebenek, S. Steckmann, S. von Stengel, W. Kemmler Osteoporoseforschungszentrum, Institut für Medizinische Physik, Universität at Erlangen-Nürnberg, Germany, Measurement of Angles in Time-of-Flight Data for the Automatic Supervision of Training Exercises.