

# A New Human Motion Analysis System Using Biomechanics 3D models

F. J. Perales, J. M. Buades, R. Mas, X. Varona, M. Gonzalez (UIB),  
A. Suescun, I. Aguinaga (CEIT)  
M. Foursa (Fraunhofer Institut)

G. Zissis (Systema) M. Touman (Synkronix)  
R. Mendoza (Kursaal Productions)  
e-mail: [paco.perales@uib.es](mailto:paco.perales@uib.es), [asuescun@ceit.es](mailto:asuescun@ceit.es)  
HUMODAN Project, UE

## Abstract

We define a process to adjust the humanoid to the morphology of the person. It could be very laborious and subjective if done manually or by selection of points, but in this poster we present a global human motion system capturing, modeling and matching as automatic process between the real person and the modeled humanoid or synthetic avatar. Automatic means that the computers propose the best matching from previous frame. Is really a new approach and is possible to applied in indoor environments or outdoor if we can control the illumination.

We have some interesting video demo of the proposed system at <http://dmi.uib.es/research/GV/Siggraph2004/index.htm> and we can offer to present at SIGGRAPH Conference and interactive demo to show our more recently results.

## 1. Introduction

We will divide the general process of the system into four stages: in the first one, we capture images of the person from different points of view and of the background, that means that a initialization process is needed to know the exact anthropometrical data of the person that is moving in front of the cameras. In the second stage, we select the humanoid with similar characteristics to the original individual, so the result of this initialization process is an avatar with the same segments length of the real person. This process is need because in high-level sport activities the accuracy is very critical (in computer vision tracking the precision or accuracy requirements can be reduced). In the following stage we apply an automatic process to obtain the humanoid adjusted to the person's measurements. The final goal is to reach an automatic recognition process, but this is a challenging task, that we are solving with less accuracy and under controlled environments. The matching criteria propose a future pose of human segments based in previous frames and the user select the best one. The last stage combines the captured images and the generated humanoid to verify the result of the process and study the values we are interested in.

## 2. The H-Anim humanoid model, Image color segmentation & Matching criteria

In the module of human definition, there are various specifications [2] for humanoids; we have chosen the one created by the group h-anim<sup>1</sup> in VRML format for its portability and adaptability to different applications. The humanoid is composed of a collection of joints and segments structured in the form of a tree. Each joint corresponds to an anatomical joint (knee, shoulder, vertebrae). Each joint has associated with it a segment, it represents, for example, the elbow and the upper arm as its segment. Image segmentation is the first step in data extraction for computer vision systems. Achieving good segmentation has turned out to be extremely difficult, and it is a complex process. Moreover, it depends on the technique used to detect the uniformity of the characteristics founded between image pixels and to isolate regions of the image that have that uniformity. Multiple techniques have been developed to achieve this goal, such as contour detection, split and merging regions, histogram thresholding, clustering, etc. A Survey can be found in [7]. In color image processing, pixel color is usually determined by three values corresponding to R (red), G (green) and B (blue). The distinctive color sets have been used with different goals, and specific sets have even been designed to be used with specific segmentation techniques.

We define a color image as a scalar function  $g = (g^1, g^2, g^3)$ , defined over image domain  $\Omega \subseteq \mathbb{R}^2$  (normally a rectangle), in such a way that  $g: \Omega \rightarrow \mathbb{R}^3$ . The image will be defined for three channels, under the hypothesis that they are good indicators of autosimilarity of regions. A segmentation of image  $g$  will be a partition of the rectangle in a finite number of regions; each one corresponding to a region of the image where components of  $g$  are approximately constant. As we will try to explicitly compute the region boundaries and of course control both their regularity and localization, we will use the principles established in [1] to define a good segmentation. To achieve our goals we consider the functional defined by Mumford-Shah (to segment gray level images) which is expressed as:

$$E(u, B) = \int_{\Omega} \sum_{i=1}^3 (u^i - g^i)^2 d\mu + \lambda \ell(B) \quad (1)$$

where  $B$  is the set of boundaries of homogenous regions that define a segmentation and  $u$  (each  $u^k$ ) is a mean value, or more generally a regularized version of  $g$  (of each  $g^k$ ) in the interior of such areas. The scale parameter  $\lambda$  in the functional (1) can be interpreted as a measure of the amount of boundary contained in the final segmentation  $B$ : if  $\lambda$  is small, we allow for many boundaries in  $B$ , if  $\lambda$  is large we allow for few boundaries.

A segmentation  $B$  of a color image  $g$  will be a finite set of piecewise affine curves - that is, finite length curves - in such a way that for each set of curves  $B$ , we are going to consider the corresponding  $u$  to be completely defined because the value of each  $u^i$  coordinate over each connected component of  $\Omega \setminus B$  is equal to the mean value of  $g^i$  in this connected component. Unless stated otherwise, we shall assume that only one  $u$  is associated with each  $B$ . Therefore, we shall write in this case  $E(B)$  instead of  $E(u, B)$ .

In the proposed system, (see poster file) the matching process is the kernel. Our main objective is to find a one-to-one correspondence between real and synthetic human body segments and joints, in every frame, in 3D space. The humanoid matching is currently done automatically; in indoor environment and, for the moment we have reached promising results in cases where occlusions are not so long in time. As we have mentioned above, the system can estimate a posture from previous frames matching in time  $t_{i-1}$ . This estimation is based on a function that evaluates visual parameters (contours, regions, color, etc.) and biomechanical conditions. The general method of this problem is known as *analysis-by-synthesis*. As we know, to search quickly in a 43-dimensional state is extremely difficult; therefore we propose some set of conditions to reduce our space search.

## 3. Conclusions

In this poster we have presented a whole system to analyzes and synthesize human movements. The main outstanding research and contributions are: a) Digital standard and portable low cost capturing system, b) Biomechanical compliant human modeling, c) Robust color segmentation process using solid mathematical background theory, d) Automatic matching process based on rules from images features and biomechanical conditions e) Genetic algorithms are used in optimization problem.

## 4. Acknowledges

Work subsidized by HUMODAN IST-2001-32202CICYT TIC2001-0931 and TIC2002-10743-E

<sup>1</sup> [www.hanim.org](http://www.hanim.org)