3D MOTION CAPTURE USING NORMAL WEBCAM

A Project Report

Submitted by:

AMARNATH KUMAR (18105128030)

PRITI GUPTA (18105128033)

MONU KUMAR RAM (18105128041)

URMILA KUMARI (18105128010)

in partial fulfilment for the award of the degree

of

BACHELOR OF TECHONOLOGY

IN

COMPUTER SCIENCE AND ENGINEERING

at

B. P. MANDAL COLLEGE OF ENGINEERING, MADHEPURA (852128)

(AFFILIATED TO ARYABHATTA KNOWLEDGE UNIVERSITY, PATNA, BIHAR (INDIA)

CERTIFICATE

This is to certify that the project titled "**3D MOTION CAPTURE USING NORMAL WEBCAM**" is the Bonafede work carried out by **AMARNATH KUMAR (18105128030), PRITI GUPTA (18105128033), MONU KUMAR RAM (18105128041), URMILA KUMARI (18105128010)** students of Bachelor of Technology in COMPUTER SCIENCE AND ENGINEERING of B. P. MANDAL COLLEG OF ENGINEERING, MADHEPURA, BIHAR (852128) affiliated to Aryabhatta Knowledge University, Patna, Bihar (INDIA) during the academic year 2018-22, in partial fulfilment of the requirements for the award of the degree of Bachelor of Technology (Computer Science and Engineering) and that the project has not formed the basis for the award previously of any other degree, diploma, fellowship or any other similar title.

Place : MADHEPURA Signature of guide

Date :

ACKNOWLEDGEMENT

This is to acknowledge all those without whom this project would not have been reality. Firstly, we would like to thanks our Project coordinator **Amarnath Kumar** for allowing us to make projects on the topic of our interest. Secondly, we would wish to thank our Project guide **Prof. Ramesh Kumar** who gave his immense support, dedicated his/her time towards it and made us understand how to make project a better one by adding more and more features to it. He guided us to make our project more interactive by adding UI designs. Without his guidance, the project would not have been complete. We wish to thank our parents as well for their undivided support and interest who inspired us and encouraged us to go our own way, without whom we would be unable to complete my project. Now, We would like to thank everyone who is involved in the project directly or indirectly. Above all, we would like to thank the Great Almighty for always having his blessing on us.

ABSTRACT

This research project is carried out to determine some of the basic human motion detection algorithm that had been founded or developed or even researched previously. This thesis report would bring a presentation of these algorithms for researchers to get a basic idea of performing an algorithm for human motion detection systems. The main algorithm being discussed here are those implementing image subtraction methods and foreground-background segmentation approach. The thesis report also is aimed to give readers a main idea of the architecture of a human motion detection system in applications. This report is also written with the purpose of documenting the design and development of a prototype human motion detection system. Here, we presented some basic ways to perform a human motion detection algorithm and also a new way to consider for background updating using spatial information instead of temporal. The experiments carried out to evaluate the performance of the prototype system is attempted and its results being recorded in this paper as well. As a conclusion, this paper is aimed to researchers interested to research on the basic idea of human motion detection algorithm using image subtraction and foreground-background segmentation techniques.

In this paper first thing that would be cleared is that Mocap is not new technology it is used since 1872 when Edward Muybridge performs Flying Horse experiment to know that if a horse ever had all four feet off the ground while trotting? So Muybridge placed cameras to capture movements of running horse and takes multiple pictures of horse and proved that statement true. After that Etienne-Jules Marey became the First person to analyse human and animal motion with video. After all these main-frame motion capture started when in 1915 Rotoscoping which is described in this paper later comes in animation techniques and it changed whole meaning of animation. Then process of basic motion capture and some techniques used i.e. how motion or movements of an actor are captured using various markers, sensors, cameras and mechanical or magnetic suits and then how these recorded data is converted and applied on a virtual actor to perform same movements. Then some applications like films, animation, medical etc. are discussed and at last a brief about some pros and cons of Mocap is stated.so overall in this paper we tried to give basic knowledge on mocap so that a non- technical or normal person can also understand that how mocap is started and how it is useful or popular now days.

INTRODUCTION

"Motion Capture" is the term used to describe the process of recording human movement and translating that movement onto a digital model. It is used in military, entertainment, sports, medical applications for validation of computer vision and robotics. In film making it refers to recording the actions of human actors, and using that information to animate digital character models in 2D or 3D computer animation. When it includes face, fingers and captures subtle expressions, it is often referred to as performance capture in motion capture sessions.

Movements of one or more actors are sampled many times per second, although with most techniques motion capture records only the movement of actors , not his or her visual appearance this animation data is mapped to a 3D model so that the model performs the same actions as the actor. This is comparable to the older technique of rotoscope such as 1978 "The Lord Of Rings" animated film where visual appearance of the motion of an actor was filmed, then the film is used as guide for the frame by frame motion of the hand-drawn animated character.

Camera movements can also be motion captured so that a virtual camera in the scene will pan, tilt, or dolly around the stage driven by a camera operator while the actor is performing, and the motion capture system can capture the camera and props as well as the actor's performance. This allows the computer-generated characters, images and sets to have the same perspective as the video images from the camera. A computer processes the data and displays the movements of the actor, providing the desired camera positions in terms of objects in the set. Retroactively obtaining camera movement data from the captured footage is known as match moving or camera tracking.

Human motion capture is the process of measuring and recording human body in a computer usable form. Interest in human motion analysis - often called “mocap”- has been growing in recent years as many potential applications have emerged (human-computer interfaces, medical applications, animation, interaction with virtual environments, video surveillance, games, etc.). This work focuses on enhancing interaction in virtual environments by reproducing user gestures directly in a 3D avatar in real-time (Horain, et al., 2005). Traditional motion capture involves special sensors (e.g. data gloves, magnetic sensors and mechanical exoskeletons) or optical markers on the performer’s body and limbs. Motion data is derived from the positions or angles of markers relative to the sensors. However, the cost and complexity of this equipment (personnel required, physical environment, etc.) is prohibitive for the general public and for many target applications. Computer vision based techniques offer an interesting alternative because they only require images from one or more cameras. As special and expensive equipment is not required, motion capture by vision is potentially a practical and inexpensive solution. We are interested in estimating 3D human motion from monocular images; this enormously increases the range of possible applications as many personal computers include a webcam. We address the problem of 3D human motion capture in real-time without markers from monocular images obtained from a webcam. We focus on capturing the motion of the upper part of the human body as our objective is to achieve more natural interactions between users and 3D virtual collaborative environments. A prototype for 3D motion capture by mesoscopic vision and virtual rendering was previously proposed in the works of Horain (Horain, et al., 2002) and Marques Soares (Marques Soares, et al., 2004). This approach consists basically of registering a 3D human upper-body model to video sequences. However, because it is designed to work in real time on a personal computer, its robustness and accuracy are currently limited and need to be improved. 3D motion capture from monocular images remains challenging open problem as many difficulties are involved; for example, the ambiguities of monocular images, partial occlusions of human body parts (e.g. crossed arms), the large number of degrees of freedom of the human body, variations in body proportions and clothing, cluttered and complex environments, image noise, etc. Moreover, the computations needed to track the human motion in the images can be very expensive, making it difficult to achieve robust real-time performance. In this thesis, new algorithms are proposed to improve the results of the motion capture methods of (Marques Soares, et al., 2004). In the following subsections, we describe our target application and some other potential applications of the work; then we discuss the main challenges and difficulties involved and present the contributions of this work.

HISTORY OF MOTION CAPTURE

The use of motion capture to animate characters on computers is relatively recent, it started in the 1970‟s and now just beginning to spread . Motion Capture is recording the movements of human body for immediate analysis. The captured information can be as simple as catching the body position in space or as forms like bvh, bip, fbx etc. Which can be used to animate 3D characters in 3D„s max, maya etc. Motion capture for animation is the superposition of human movement on their virtual identities this capture can e direct such as the animation of virtual function of movement of an arm or indirect such as that of human hand with a more thorough as the effect of light color. To make the most convincing human movement in “snow white”, Disney studios design an animation film on a film or real players.

BVH (Bio-Vision hierarchy)

The Biovision Hierarchy (BVH) character animation file format was developed by Biovision, a defunct motion capture services company, to give motion capture data to customers. This format largely displaced an earlier format Biovision providing skeleton hierarchy information as well as motion data.

As of 2019, BVH is widely used and most 3D applications support importing and exporting files in this format.

Examples of software using files in BVH format:

1. Lightwave 3D
2. 3ds Max (version 9 or later)
3. Blender
4. Clara.io
5. Cobalt
6. Daz Studio
7. Esenthel Engine
8. Lifeforms
9. MakeHuman

10.Maya

11.Modo

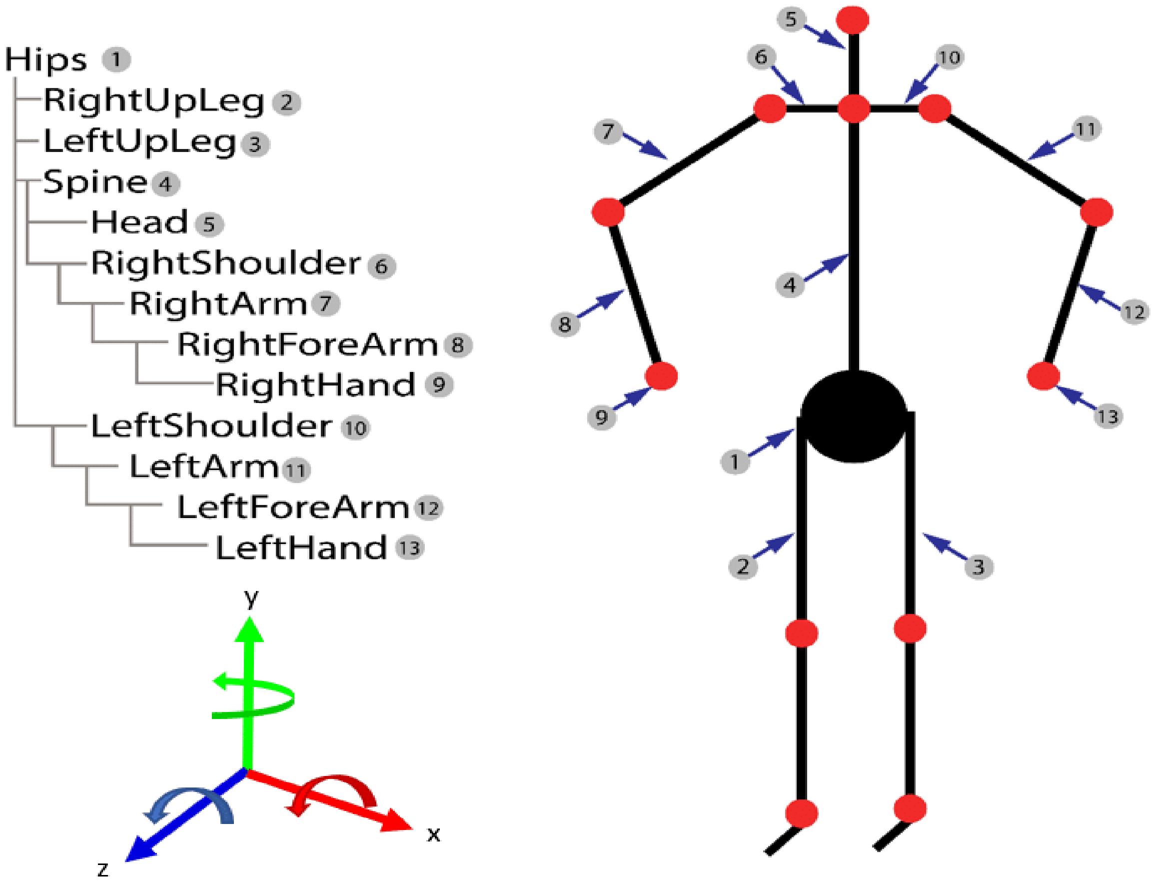
12.Poser

13.Seamless3d Free open source modeller

14.Second Life / OpenSim

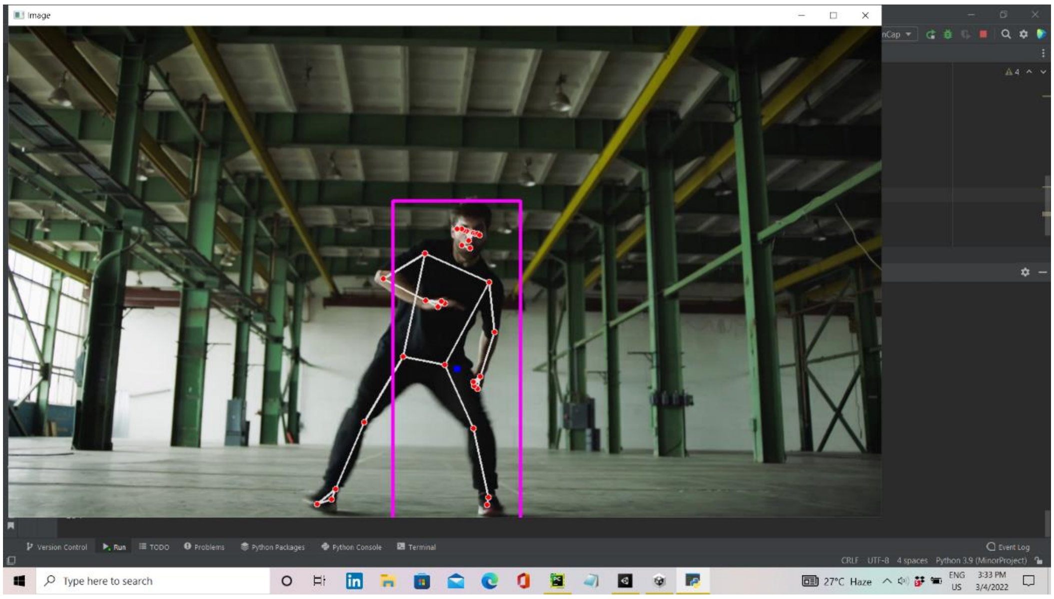
15.Avimator and its Qt port Qavimator

16.Maxon Cinema 4d



FBX (Film Box)

An FBX (.fbx) file is a format used to exchange 3D geometry and animation data. You can use different programs to open, edit, and export high-fidelity 2D and 3D files. FBX files are used in film, game, and Augmented Reality and Virtual Reality (AR/VR) development.



DIFFERENT TYPES OF MOTION CAPTURE

Motion Capture technology can be achieved by using the following three types of techniques:

1. Mechanical motion capture
2. Optical motion capture
3. Magnetic motion capture

Now although this technique is effective, it still contains some problems (weight, Cost). But against any doubt that the motion capture will become one of the basic tools of animation.

Mechanical Motion capture:

This technique of motion capture is achieved through the use of an exoskeleton. Each joint is then connected to an angular encoder. The value of movement of each encoder (rotation etc...) is recorded by a computer that by knowing the relative position encoders (and therefore joints) can rebuild these movements on the screen using software. An offset is applied to each encoder. because it is very difficult to match exactly their position with that of the real relationship (and especially in the case of human movements).

Fig 3.1 Mechanical Motion Capture using Exo skeleton

Advantages and Disadvantages

1. This technique offers high precision and it has the advantage of not being influenced by external factors (such as quality or the number of cameras for Optical MOCAP).

2. But the catch is limited by mechanical constraints related to the implementation of the encoders and the exoskeleton. It should be noted that the exoskeleton generally use wired connections to connect the encoders to the computer. For example, there is much more difficult to move with a fairly heavy exoskeleton and connected to a large number of simple son with small reflective sphere. The freedom of movement is rather limited.

3. The accuracy of reproduction of the movement depends on the position encoders and modelling of the skeleton. It must match the size of the exoskeleton at each morphology. The big disadvantage comes from the coders themselves because if they are of great precision between them it cannot move the object to capture in a so true. In effect, then use the method of optical positioning to place the animation in a decor. Finally, each object to animate to need an exoskeleton over it is quite complicated to measure the interaction of several exoskeleton. Thereby bringing about a scene involving several people will be very difficult to implement.

Magnetic motion capture:

Magnetic motion capture is done through a field of electro-Magenta is introduced in which sensors are coils of sensors electrocutes, Les son are represented on a place mark in 3 axes x, y, z. To determine their position on the capture field disturbance created by a son through an antenna then we can know its orientation.

Fig 3.2 Magnetic Field Transmitter Source

Advantages and disadvantages

1. The advantage of this method is that data captured is accurate and no further calculations excluding from the calculation of position is useful in handling.
2. But any metal object disturbs the magnetic field and distorts the data.

Optical Motion Capture:

The capture is based on optical shooting several synchronized cameras, the synthesis of coordinates (x, y) of the same object from different angles allows to deduce the coordinates (x, y, z). This method involves the consideration of complex problems such as optical parallax. distortion lens used, etc. The signal thus undergoes many interpolations. However. a correct calibration of these parameters will help in high accuracy of data collected. To determine their position on the capture field disturbance created by a son through an antenna then we can know its orientation.

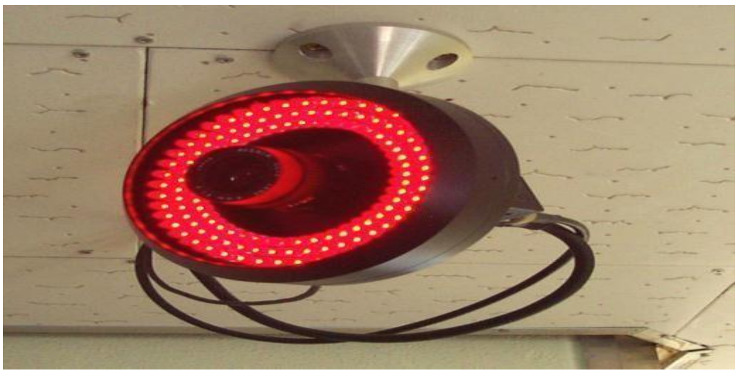


Fig 3.3 Camera emitting Infrared Radiations

The operating principle is similar to radar: the cameras emit radiation usually infrared, reflected by the markers and then returned to the same cameras. Checking the information of each camera (minimum two cameras) to determine the position of markers in virtual space.

MAIN CHALENGES

We focus on 3D motion capture by monocular vision in real-time without markers (Horain, et al., 2002). This computer vision problem is inherently difficult for because of insufficient information from images, ambiguities in the projection of monocular images, the large number of human pose parameters to estimate and self-occlusion of body parts. In this section, we describe the main challenges addressed in this work.

Lack of depth information: inferring the 3D pose from images obtained with only one camera is an ill-posed problem. Forward and backward movements with respect to the camera‟s viewpoint (depth direction) lead to more than one possible solution due to the ambiguities in monocular images. These ambiguities result in motion-mistracking.

High-dimensional search: in order to estimate the 3D human pose, several works (Sminchisescu, et al., 2002), (Delamarre, et al., 2001) use a 3D human model to infer the correct joint angle values (degrees of freedom) of each body part (head, neck, arms, forearms, 25 hands, legs, etc.). Some other works (Ramanan, et al., 2003), (Noriega, et al., 2007) try to infer the position of each body part separately before enforcing joint connectivity. In both cases, the number of parameters to be estimated is very high (between 20 and 60 dimensions); therefore, finding the optimal pose is a search problem in a high dimensional space. This requires intensive computation that is difficult to achieve real-time.

Occlusions of human body parts: tracking 3D human motion from monocular images can become a difficult task as there are poses where some body parts are occluded by others (e.g. walking, arms crossed). In order to cope with occlusions, several works (Ning, et al., 2004), (Sidenbladh, et al., 2002) use a learned motion model or exploit a temporal coherence in order to track the pose through body parts occlusions.

Clothing variations: different clothing in humans can be a problem (e.g. clothing of different shapes, or with color, edges and complex textures) causing inaccurate observations leading to ambiguities and consequently, incorrect 3D pose estimates.

Variations in human body proportions: many 3D pose estimation techniques (Deutscher, et al., 2000), (Sminchisescu, et al., 2002) use a 3D computer model of the human that must be matched to the actor in the image. Such model usually has predefined body part sizes and shapes. Thus, if the human model has different proportions from the real subject, the matching between the model projections and the actor cannot be accurate and failures may occur. For example, if the arms of the model are larger than the arms of the subject, the model arms can intersect erroneously with other body parts.

General motions: several works on 3D motion capture (Urtasun, et al., 2004) use learned motion models to track cyclic or repetitive motions (e.g. walking, running, golf swing) however, general motions are difficult to track as they are highly variable and unpredictable, making the learned model useless. Several works address the problem of tracking general motions by propagating multiple hypotheses at each time (e.g. particle filter approaches), this often provides robust and accurate results, however, computation times can be very high as a large number of hypotheses must be propagated.

Complex environments: lighting changes and cluttered or dynamic backgrounds (e.g. public spaces) from video sequences can be a problem as imprecise observations (e.g. clutter edges, segmentation failures) can make the pose estimation procedure fails. Lighting changes can cause misclassifications in color segmentation or the emergence of new segments in the scene. Shadows, textures and objects in cluttered backgrounds can produce distracting image measurements (e.g. color, edges, optical flow).

Motion capture technologies

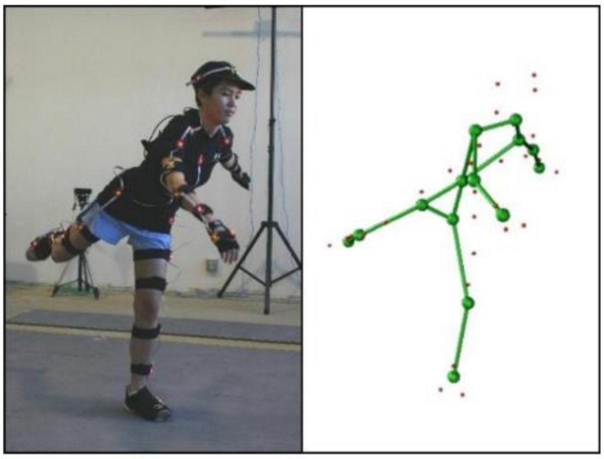
Mocap technologies generally include synchronized cameras and special suits with markers or sensors worn by the performers. Markers are set at body parts (or joints) and are tracked in order to identify the motion by their positions or angles. Only the motion of the actors, not their visual appearance, is recorded and this motion is the mapped to a 3D model by computer.

Motion capture techniques can be classified by their input methods, namely optical, mechanical, magnetic, acoustic and inertial. Each of these inputs (or a combination of them) is tracked, ideally at least twice the frequency of the desired motion.

Optical systems: In optical motion capture systems, data acquisition is implemented using special markers attached to the actor (Opti track, 2010); this approach uses at least three synchronized cameras and proper lighting to estimate the performer's position in 3D space. These systems produce data with 3D locations for each marker (Figure 2-1). Two types of markers can be used: passive and active.

Passive markers: Passive optical system use markers coated with a retro-reflective material to reflect back light that is generated near the cameras lens. The operating principle is similar to radar: the cameras emit radiation (usually infrared), which is reflected by the markers and returned to the same camera. The cameras are sensitive to a narrow band of wavelengths and perceive the markers as bright spots. Such a system typically consists of 6 to 24 cameras. Passive systems do not require the user to wear wires or electronic equipment. The markers are usually attached directly to the skin, or to a full body lycra suit designed for motion capture (Qualisys, 2010).

Active markers: Rather than reflecting back externally generated light, the markers themselves are powered to emit their own infrared light. Active optical systems triangulate positions by illuminating one LED at a time very quickly or multiple LEDs with software to identify them by their relative positions. Some systems can modulate the amplitude or pulse width in order to provide marker ID. This unique marker ID provides much cleaner data than passive marker systems (PhaseSpace, 2010).



:Optical motion capture system. Actor wearing an optical motion capture suit with infrared markers (left). 3D human pose inferred from the relative positions of each marker (right) (PhaseSpace, 2010).

Inertial systems: Inertial systems are based on miniature inertial sensors (gyroscopes and accelerometers) with sensor fusion algorithms to measure the rotational rates. Motion data is transmitted wirelessly to a virtual skeleton in software. Basically, these systems are similar to the Nintendo Wii controllers but are more sensitive and more accurate. Inertial systems are practical because they are portable and do not require large capture areas or complicated calibration (IGS-190-M, 2010). A disadvantage is that the positional drift that can compound in time if the setup is not correct due to poor calibration.

Computer vision based systems

Using computer vision techniques to acquires the human body motion is one of the most attractive and practical solutions as it does not require any expensive or invasive hardware or markers (only cameras are required) and it can work outdoors (in streets, offices, parks). Algorithms have been proposed that capture human motion at near real-time frame rates; however, they mostly rely on multi camera systems under controlled conditions, which limit their applicability. Some monocular vision approaches (Agarwal, et al., 2006), (Urtasun, et al., 2006) aim at capturing specific motions (walking, golf swinging, jumping, etc.) using some learning model or tracking motion for certain parts of the body. Some other systems can track unconstrained motion, but do not run in real-time (Sminchisescu, et al., 2003). Recently, 3D image sensors have been used to capture the 3D body shape and disambiguate poses using depth measurements. Time-of-flight sensors (Ganapathi, et al., 2010), (Kolb, et al., 2010) and active triangulation system (e.g.

Microsoft Kinect (Kinect, 2010)) are such dedicated sensors. Real-time markerless tracking of human pose by monocular computer vision remains a hard yet relevant problem

Human motion capture by computer vision

Human pose estimation from images is an unsolved and currently active field of research with significant scientific and computational challenge (Sminchisescu, 2007). As this is the main topic of this work, we present a more detailed analysis of computer vision-based techniques. We consider both real-time and non-real-time systems, as well as single and multi-camera systems.

Systems for motion capture using computer vision are first discussed based on the image descriptor or features that they use. Human pose can then be estimated using a generative approach or a discriminative approach (Sminchisescu, 2007). Tracking the pose between consecutive frames of a video sequence can ensure the temporal coherence of the human motion. Finally learned motion models can be used to enhance the robustness of the motion tracking.

Image features for motion capture

In order to estimate the 3D pose, image features are extracted from input images. A feature or image descriptor can be defined as a piece of low-level visual information extracted from an image to solve a specified task. Image descriptors are generally extracted using probability distributions (e.g. color histograms), neighbourhood operations (e.g. edges, optical flow) or thresholding the pixel values (e.g. foreground or colour classification). The choice of features depends greatly on the problem to be solved. In human motion capture, the image descriptors are generally used as cues to find the position of each body part and thus to estimate the full 3D human pose. Image descriptors commonly used in the literature include color, silhouettes, edges and motion (Poppe, 2007), (Moeslund, et al.).

Generative approaches

These approaches estimate the human pose using a prior model of the human body, parameterized by the kinematic tree of the articulations and the body dimensions. The pose of the human body model is described by a vector of parameters.

Generative approaches differ essentially in the manner in which data is associated with the 3D model. In this case, we can identify two methods: reconstruction-based and appearance-based. The first method (reconstruction-based) try to fit the 3D model to a 3D cloud obtained from multiple cameras (Urtasun, et al., 2004), (Ziegler, et al.). In the appearance-based techniques, a human model with a defined pose is projected into the input image and several features (section 2.3.1) are extracted from the image (Lee, et al., 2002), (Ramanan, et al.), (Sminchisescu, et al., 2003). In both techniques, the human pose is estimated by finding the vector of model parameters that best fits the input data.

Finding the pose that best matches the image features can be a very difficult task, because occlusions of body parts may occur and same image feature can match different 3D poses (ambiguities). Several works propose different human body models and methods to estimate and track the human pose over time. Some learning methods are also used to improve the motion capture results.

Discriminative approaches

Discriminative approaches do not use an explicit human body model based representation. Instead, they infer the human pose directly from the image observations or features, using training examples to establish a direct relationship (or mapping) between the image observations and human poses; this assumes that the set of typical human poses is far smaller than the set of kinematically possible ones. Therefore, the training data needs to generalize well to observed variations over body configuration, body dimensions, viewpoint and appearance. The mapping method must also account for the highly non-linearity of the mapping between the feature and the pose space.

These approaches have the advantage of avoiding the need for explicit initialization and accurate 3D modeling and rendering. They can also be used to initialize generative approaches as in (Fossati, et al., 2007). However, they require a sufficient number of training examples in order to infer human pose properly, and they tend to be more sensitive to background clutter than generative approaches because an explicit model is not available for background masking.

Pose tracking

Pose tracking is the process of following and estimating the human pose from frame to frame in a video sequence. Tracking the human pose has several advantages:

a) The difficult task of searching the high-dimensional pose space is alleviated as pose differences between frames are usually small,
b) Temporal coherence is ensured when dealing with image projection ambiguities,
c) The complexity of the pose estimation is reduced as an initial pose estimate is provided at each frame.

Generally, there are two strategies for tracking the pose: those that maintain or predict only one hypothesis (pose configuration) at each frame (single hypothesis tracking) and, those that propagate several hypotheses (multiple hypothesis tracking) or solutions per frame.

Dynamic models

Dynamic models can be used to encode the expected dynamics of a human motion, e.g. periodic motions such as walking, running, swinging, etc. They are used as predictive priors for tracking, providing more stable tracking at reduced computational cost. They are often learned from training data (e.g. body pose parameters) acquired with a motion capture system.

Using such models, the robustness of tracking can be enhanced even with incomplete information or occlusions, because the prior motion model allows spurious and distracting information to be discarded. They can also recover from tracking failures by using the motion priors as new starting points. Nevertheless, dynamic models have the disadvantage of depending significantly on the scope of the available training data; the set of exemplars must be sufficiently large to account for any variation that may occur in the captured movement. Using a strong motion prior limits the tracking essentially to the set of actions learnt beforehand.

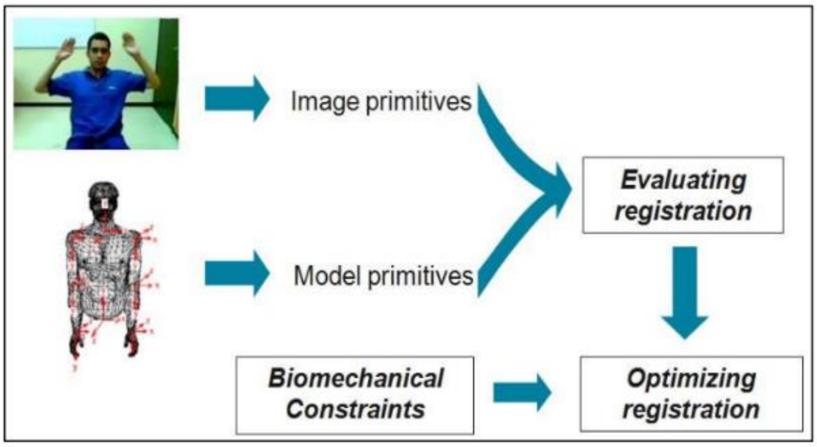
Dynamic models for motion capture can be divided into two classes: high-dimensional models that are learned directly from the original pose space and low-dimensional models that consist in a reduced latent space with lower dimension, in which tracking is performed.

Our baseline approach for 3D motion capture

We describe our baseline approach previously proposed in the works of (Marques Soares, et al., 2004). Basically, our approach for 3D motion capture from real-time 49 monocular vision consists of registering a 3D articulated model of the upper human body to 2D video sequences. No learning is involved in this algorithm because we aim to capture general human gestures. Our baseline approach works with the following assumptions:

1. The user remains seated while making gestures in front of his computer.
2. The shirt of the user has short sleeves and a uniform color.
3. The background is static.
4. In the first image of a video sequence, the user will be assumed in a fronto-parallel pose with no body part occlusions.

As seen in Figure 2-13, we extract primitives from the input image and from our human model. The similarity between the image and model primitives is then evaluated using a matching cost function. Our registration process consists searches for the pose of the 3D body model that optimally matches the primitives extracted from the 2D image. Biomechanical constraints allow poses that physically cannot be reached by the human body to be invalidated (Marques Soares, et al., 2004). At each new frame of the video sequence, 3D/2D registration is done starting from the configuration resulting from the previous frame. The following sections explain our approach in detail.



The general strategy of our approach for 3D motion capture

FUTURE SCOPE

Over the past several years, the adoption of cloud technology has completely changed the way we work and has allowed many sectors in the entertainment industry to take their offerings to the next level. From video games and movies to broadcast entertainment, many sectors have already undergone a serious and accelerated digital transformation thanks to the advantages of cloud computing and the benefits are seemingly endless - scalability, agility, cost-saving, the list goes on.

Motion capture is one of the first technologies that saw a period of rapid growth and expansion which led it to become an integral part of many different industries, ranging from sports science to medical applications and entertainment. Even now, the mocap space is continually evolving and is looking to expand into virtual and augmented realities. However, the mocap industry is only now ready to tap into the potential of the cloud - but what would that even look like?

Being able to upload motion capture data directly to the cloud and process it, without the need for powerful on-premises hardware, is a significant step forward for the motion capture industry. Traditionally, the more motion capture data tracked for processing, the more processing power required for your PC - but not with the cloud. And this has far-reaching implications for many industries, not just entertainment. Those working in sports science, research, and ergonomics - which often lack the compute power of large creative studios - can now process large samples of data quickly. A whole football team could be simultaneously captured and processed in the cloud effectively and quickly, something that would have required powerful computers before the availability of cloud technology.

Not only would the cloud allow for fast processing speeds, but it would also facilitate global collaboration between users in entirely different locations. For example, a studio recording full-body motion data in the US can stream, store, and process that data directly into the cloud, with a corresponding animation team in Singapore able to view and download specific data to drive their animated scenes. This would hugely speed up the collaboration time between teams, allowing almost instant access to the data for multiple users, and increase the efficiency of remote working.

Despite an industry shift towards a more collaborative approach to working already underway, 2020 set a clear path for the future of remote working. Millions of people around the world had to quickly adapt to home-based working setups, with large creative studios temporarily losing access to their large physical spaces. It's clear that hybrid, and in some cases entirely remote setups, are here to stay, and during the pandemic, some of our customers have been creative by taking their suits home to record full-body movements to drive professional animated works – this inspired us to launch our own cloud computing offering, Motion Cloud. Cloud access can facilitate remote working in more intuitive ways, making things more straightforward than ever before.

This is only the beginning of the motion capture industry's potential use of the cloud and we cannot wait to see what innovative solutions and workflows appear in the coming years.

CONCLUSION

Although the motion capture requires some technical means, we can quite get what to do it yourself at home in a reasonable cost that can make your own short film. Motion capture is a major in the field of cinemas as you can reprocess the image in a more simple in fact it is easier to modify an image captured a classic scene, all although this is too expensive, but it is also a major asset in medicine, for example it can be used to measure the benefit of a transaction via recording of the movement of patient before or after the operation (such as in the case of application prosthesis or simply at a medical classic in the future perhaps).

BIBLIOGRAPHY:

https://www.fogscreen.com
https://wikipedia.com
https://www.engadget.com/tag/interface/
https://www.broadcastnow.co.uk/
https://www.youtube.com/