Creation of digital risk doubles using motion capture and photogrammetry for computer-generated imagery content

G J Silva-Bolivar¹, C A Reyes Muñoz¹, S Londoño-Marín¹, and J J Castro-Maldonado¹
¹ Grupo de Investigación en Gerencia y Aplicación de la Ciencia y la Tecnología, Servicio Nacional de Aprendizaje, Medellín, Colombia

E-mail: gsilva42@misena.edu.co

Abstract. In Colombia, the audiovisual industry has grown exponentially thanks to the policies on the orange economy, which seeks to impact the level of development in the artistic, cultural, and information and communication technologies sectors. In order to realize as a power in this sector worldwide, Colombia advances rapidly in the strengthening of the industry, venturing into the implementation of new technologies in special effects and post-production processes. In the flow of current film production, digital doubles, digital clones of the actor that allow dubbing for risky scenes, characters with particular aesthetics or actors that cannot be present in the set. The film industry has adopted in order to reduce the risks in the times of work of the actors, the reduction of the risks in the scenes of action and the decrease of the costs in the make-up and the utility. This article shows the process of developing a digital double in computer-generated imagery techniques, and how to achieve that digital realism that allows to pass, unnoticed dubbing, in order to socialize the industry knowledge and explore this type of innovations in their processes, following a sequential methodology from its conception to its conception.

1. Introduction
La Photogrammetry consists in the geometric reconstruction of an object from photographs, therefore, initially it is posed as a geometric problem of reconstruction of homologous rays (which go to the same point) from two different points of view.

According to Figure 1, given a beam of perspective rays $\Gamma$ and a set of half-lines $SA, SB, SC$. from an $S$ point of view, the reconstruction of points $A, B, C$ of the object $\Sigma$ can only be carried Out geometrically knowing the distances of these points. Obviously, it is not the case of photogrammetry, since it cannot measure these distances, only images are available. In this case it can only be done by intersecting homologous rays from another point of view (in this case, two frames) [1,2].

This technique has been used in areas such as cartography, topography and / or geology [3], however, thanks to the advancement of optical sensor technology and photography equipment this technique has expanded to other fields [4–6] as is the case of creative industries for the development of audiovisual content.

The creative industries that are part of the orange economy have evolved in structure, business models and the relationship between creators and users, giving rise to new cultural expressions [7]. According to John Howkins “The creative economy, includes the sectors in which the value of their goods and services is based on intellectual property: architecture, visual and performing arts, crafts, cinema, design, publishing, research and development, games and toys, fashion, music, advertising,
software, TV and radio, and video games” [8]. Colombia advances rapidly in strengthening the industry, venturing into the implementation of new technologies in special effects and Postproduction [9].

![Figure 1. Fundamental problem of photogrammetry [1].](image)

In the search to match the gap between the levels of the national industry with respect to the international scene, the investigation of procedures that help to reduce costs, optimize processes and increase production quality is essential. The search for these factors among the special effects generated by computer [10], represent an opportunity for Colombian productions, since they allow to obtain great results without major modification to the infrastructure that production studies can acquire. In the current technological resources, we find an aspect that can improve the dubbing of actors in Colombian productions through digital actors, as they do in AAA productions, through motion capture and photogrammetry [11].

2. Experimental procedure
This article aims to detail the process of developing a digital double in computer-generated imagery (CGI) techniques, using specialized equipment and software for motion capture, photogrammetry and 3D animation; the sequential phases that lead to the conception of a digital double and its insertion in an audiovisual or interactive product are shown below. In general, the process can be described in three parts. The first is the obtaining of the digital model identical to the actor, the preparation of the model for the animation process and the dubbing of the performance through the digital double.

2.1. Obtaining the digital model

2.1.1. Photogrammetry. Photogrammetry allows obtaining three-dimensional models with hyperrealistic quality [12], because the result is a 3D mesh that arises from the calculation of the iteration of common points in multiple photographs of an object, as mentioned earlier, this fidelity in the geometry adds to the adhesion of a texture with the real colors of the model for each of the points of that high density mesh. Allowing thus, the reproduction with high fidelity of the model through an automated process, which could cost a lot to an artist.

2.1.2. Process in photoscan. For the creation of photogrammetric models there are several software alternatives on the market, for the creation of this model Agisoft PhotoScan was chosen, for the fidelity of the models created, for the high density of the point clouds that can be created and for the number of photographs that can be included for production [13]. Agisoft PhotoScan is an intelligent tool that processes photographs and calculates the spatial axis of the camera that made the shot, to position and calculate the cloud of points and create a digital image from common points in the texture.

By following a linear workflow consisting of the location of the cameras, the creation of the point cloud, the creation of the high-density mesh and the application of the texture on the mesh we obtain a 3D digital representation of the exportable model in object file (OBJ), one of the recognizable standard formats for most 3D element editing software, Figure 2.
2.2. Model configuration for animation
The model obtained in the previous process has visual fidelity, but it is an imperfect mesh in its composition, which has defects in the direction of the lines and accidents produced as noise of the automatic process. To configure the model for animation it is necessary to perform a process called retopology.

2.2.1. Restoration in blender, textures in photoshop and rigging in maya. Obtaining an initial basis of the photogrammetry process, a restoration in the missing segments is required through the retopology process using blender 3D. The process consists of creating new polygons drawn on the photogrammetric mesh, to generate uniformity in the direction of the faces and the least number of residual polygons as can be seen in Figure 3 and Figure 4.

At the end of the retopology process, the UV texture map is created, to edit the texture acquired in photoscan, by a graphic artist in photoshop, as seen in Figure 5. Rigging is a process in which the model for animation is configured through the construction of a “skeleton”, which allows the geometry to be doubled when inheriting the Motion capture and performing the Retargeting. Figure 6.
2.2.2. Blender hair and eyes creation. Considering that for photogrammetric capture, hair should be ignored (it cannot be rebuilt in the software due to lack of volume). Using the blender 3D Program, body, facial and hair will be designed that will have double digital, by means of a simulation plugin. The correct setting of the values depends on the hair used as a reference. From a reference image and a sphere, the shape of the eye is achieved, importing the texture on the model, Figure 7. For the hair to have movement, the hair dynamics option is used, and a material is assigned, Figure 8.

2.3. Animation of the digital double

2.3.1. Motion capture. In the early 1980s, Tom Calvert, a professor of kinesiology and computer science, added potentiometers to a body and used the sensor's output to handle animated figures [14]. After this the first visual monitoring systems emerged, using special markers on the bodies of the actors, recognized by several cameras. In 1988, during the Siggraph conference, they introduced Mike “the talking head”, thus demonstrating for the first time a performance Real-time digital actor using MoCap [15]. Currently, MoCap is present in the development of films such as The Lord of the Rings and Avatar, in video games such as FIFA, among many entertainment products.

2.3.2. Proper mocap equipment calibration procedure. The MoCap system in this experiment is composed of 8 infrared cameras for positioning a suit with passive markers. For the calibration of the system the control of the cameras for the optical system is used: Focus, diaphragm aperture and focal length, verifying that all the markers are visible. The ArcWand2 device allows you to identify the position of the cameras with respect to the others, enumerate them and define the volume of the capture and its origin, Figure 9 and Figure 10.

2.3.3. Work process in shogun live, shogun post and retargeting. Shogun live is the software that allows MoCap in real time with the vicon system and is designed for specific work on entertainment products. Figure 11 Here the subject who is going to perform the performance is calibrated based on the personalized layout of the markers, this prevents it from being confused with another actor in the scene.

Having captured the movement of the actors, the process of cleaning the frames must be carried out to guarantee the continuity of the movements, this is done by shogun post through automatic or manual processes. The animation is exported in one of the standard formats for this Filmbox (FBX) and OBJ process.

The final process to configure the digital actor is retargeting. In 3D editing software, the following must be imported: The model achieved in photogrammetry after retopology, rigged and textured; and the animation obtained in the motion capture system after cleaning. The bone system created in the rigging of the model is related, with the positioning information contained in the MoCap animation. Now in which the points inherit their rotation and translation to the model, they will produce the desired joint displacement. Figure 12.
3. Results and discussion
At this point, a three-dimensional digital model was achieved, which looks like the real actor and an assembly with the realistic movements obtained. The digital model was inserted in a production, to make the dubbing of 3 planes, one with details of the face (Figure 13) and two general planes (Figure 14), and they were inserted into the scenes through tracking techniques for the composition digital video.

Figure 9. Calibration process.

Figure 10. Numbering of cameras.

Figure 11. Bone system.

Figure 12. Motion capture application in the digital model.

Figure 13. Comparison of double digital results.

Figure 14. Inserted map of the digital double in audiovisual production.
4. Conclusions
This article described the process and technique of creating realistic digital doubles. The origins of the technique are mentioned, and the process is explained step by step. When projecting the final short film, in all the projections, the fact that the actor had been replaced by a digital double in three planes of the video went unnoticed among the viewers.

Finally, a list of advantages and disadvantages is presented where aspects such as the cost of using this technique, or the benefits in terms of performance and quality are considered in Table 1.

Table 1. Advantages and disadvantages.

| Advantages                                                                 | Disadvantages                                                                 |
|---------------------------------------------------------------------------|-------------------------------------------------------------------------------|
| Cost savings for hiring risk doubles for dangerous scenes and hiring the  | The initial cost for the acquisition of equipment necessary for the            |
| most expensive actors in a production by reducing filming hours.          | implementation of the workflow is high.                                       |
| La Photogrammetry as a base model reduces modeling and texturing times   | Photogrammetry implies a complex learning curve to obtain clean models from   |
| for a 3D artist.                                                         | the shot. This implies an in-depth study of the standardization of              |
| Motion capture as a base animation allows you to reduce animation times  | The workflow between motion capture software is very tightly controlled by     |
| in the software.                                                         | the hardware manufacturers. This prevents customization of the workflow.      |

References
[1] Sánchez Sobrino J A 2001 *Introducción a la fotogrametría digital* (Madrid: Escuela Técnica Superior de Ingenieros de Caminos, Canales y Puertos)
[2] Thevara D J, Vasanth Kumar C H 2018 Application of photogrammetry to automated finishing operations *IOP Conference Series: Materials Science and Engineering* 402(1) 012025
[3] Paixão A, Resende R, Fortunato E 2018 Photogrammetry for digital reconstruction of railway ballast particles – A cost-efficient method *Construction and Building Materials* 191 963
[4] Candotti C T, Gelain G M, Antonioli A, Araújo L M, Vieira A, Loss J F 2019 Repeatability and reproducibility of postural variables by photogrammetry *Journal of Manipulative and Physiological Therapeutics* 42(5) 372
[5] Stuani V T, Ferreira R, Manfredi G G, Cardoso M V, Sant’Ana A C 2019 Photogrammetry as an alternative for acquiring digital dental models: A proof of concept *Medical Hypotheses* 128 43
[6] Andrade R M, Guimarães L R, Ribeiro A P, Pasqual Marques A, Crivello Jr O, Gonçalves de Carvalho B K, Amado João S M A 2019 Reliability in mandibular movement evaluation using photogrammetry in patients with temporomandibular disorders *Journal of Manipulative and Physiological Therapeutics* 42(4) 267
[7] Blanco-Valbuena C E, Bernal-Torres C A, Camacho F, Díaz-Olaya M 2018 Industrias creativas y culturales: estudio desde el enfoque de la gestión del conocimiento *Información Tecnológica* 29(3) 15
[8] Buitrago Restrepo F, Duque Márquez I 2013 *La economía naranja, una oportunidad infinita* (New York: Banco Interamericano de Desarrollo)
[9] Rubio Alcover A 2006 *La postproducción cinematográfica en la era digital: Efectos expresivos y narrativos* (Castellon: Universitat Jaume I)
[10] Le B H, Zhu M, Deng Z 2013 Marker optimization for facial motion acquisition and deformation *IEEE transactions on visualization and computer graphics* 19(11) 1859
[11] Kopniak P 2015 Motion capture using multiple kinect controllers *Przegląd Elektrotechniczny* 1(8) 26
[12] Ruiz J, Roges L, Voces A G 2015 Hyper-realistic and interactive three-dimensional visualization: Holy Chamber and Jewels of the Cathedral of Oviedo *Virtual Archaeology Review* 6(12) 69
[13] Xiaohong Z, Shengliang Z, Fengtian C 2013 The application of agisoft photoscan in UAV aerial photographic image data processing *Value Engineering* 20 230
[14] Rincon P, Zambrano E 2012 *Motion capture* (Venezuela: Universidad Simón Bolívar)
[15] Parent R Ebert D S, Gould D, Gross M Kazmier C, et al. 2010 *Computer animation complete. All-in-one: learn motion capture, characteristic, point-based, and Maya winning techniques* (Burlington: Morgan Kaufmann Publishers)