Signing Avatar Motion: Combining Naturality and Anonymity

Félix Bigand, Annelies Braffort, E. Prigent, Bastien Berret

▶ **To cite this version:**

Félix Bigand, Annelies Braffort, E. Prigent, Bastien Berret. Signing Avatar Motion: Combining Naturality and Anonymity. International Workshop on Sign Language Translation and Avatar Technology, Sep 2019, Hambourg, Germany. hal-02400929

HAL Id: hal-02400929
https://hal.science/hal-02400929

Submitted on 16 Dec 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

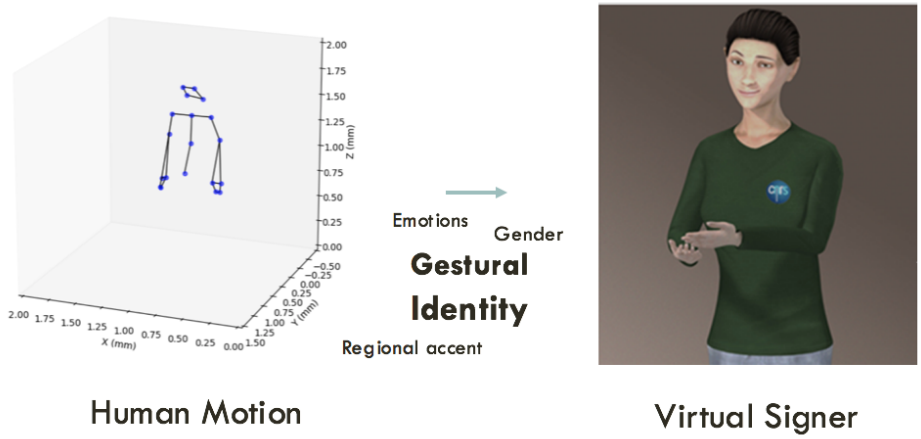
Signing Avatar Motion: Combining Naturality and Anonymity

Félix Bigand
felix.bigand@limsi.fr
LIMSI, CNRS, Paris-Sud University, Paris-Saclay University
Orsay, France

Annelies Braffort
annelies.braffort@limsi.fr
LIMSI, CNRS, Paris-Saclay University
Orsay, France

Elise Prigent
elise.prigent@limsi.fr
LIMSI, CNRS, Paris-Sud University, Paris-Saclay University
Orsay, France

Bastien Berret
bastien.berret@u-psud.fr
CIAMS, Paris-Sud University, Paris-Saclay University, IUF
Orsay, France



ABSTRACT

This paper deals with the movement of virtual signers, and more particularly with the notion of biological motion and the issue of anonymization. Animating an avatar from motion models as biological as possible is important to ensure perceptual acceptability. Recording real movements thanks to motion capture systems and mapping them to an avatar can ensure this acceptability. However, such accurate systems may convey extra-linguistic information such as the signer's identity. For some applications, it will be necessary to determine the parameters of the movement that carry identity in order to control them.

KEYWORDS

Sign Language, Virtual Signer, Signing Avatar, Animation, Biological Motion, Identity, Anonymization

1 INTRODUCTION

The use of virtual signers (or signing avatars) brings many advantages: It is possible to reuse, adapt, or modify the content of the animation; the appearance of avatars (age, gender, style of clothing, etc.) can be modified according to the target population; they can be dynamic and interactive [13]. However, their use is still limited to date.

A first reason is that the existing models of language are still very incomplete. Research on SL is recent compared to that of spoken languages. SL grammar is not yet fully described, in particular with regard to all phenomena related to the visual-gestural nature of the channel used: multilinearity of articulators. Criticisms of animations particularly concern non-manual parameters (facial expression, movement of the mouth, head and chest) [13] that are important elements of the language. In addition, we can also mention the linguistic use of space, and the linguistic structures used for depiction (classifiers). It should be noted that promising work is currently being done on these aspects [8], [7], but much remains to be done.

A second reason concerns the quality of the animations. When they are automatically generated from oversimplified models, motions will be too robotic. An alternative is to record movements on a person, for example using a motion capture system, and to animate the avatar from these pre-recorded movements. Such accurate systems provide much more natural motion that can be applied to the avatar, but they also can make the actor identifiable. The

SLTAT 2019, September 29, 2019, Hamburg, Germany

ideal would be to have motion models or processes able to hide the identity of the signer, while keeping motions as 'natural' as possible and preserving the linguistic information.

In this abstract, we explore this issue and more particularly that of the identity conveyed by motion and what this implies in the context of virtual signers.

2 BIOLOGICAL MOTION AND PERCEPTUAL ACCEPTABILITY

The properties of human motion have been studied for more than a century. It has been shown that human movements have remarkable properties, such as the 2/3 power law (implying a relationship between the speed and the radius of curvature of the trajectory), the principle of isochrony (movement duration is almost independent of its amplitude), the law of asymmetry in vertical movements due to gravity (depending on whether the movement is upward or downward), or the Fitts' law (which expresses the minimum time required to reach a target according to its width and distance).

The term 'biological motion' defines motion that comes from actions of a biological organism, that is humans or animals. The ability to perceive biological motion seems to appear very early in human life. A study conducted on infants showed that newborns are able to differentiate between biological and non-biological kinematics despite their immature visual system. The authors define biological kinematics as the movement that satisfies the constraints of the 2/3 power law mentioned above. In the study, infants spent more time observing the movements of light points that did not comply with this law. These results could be explained by the 'expectation violation model', i.e. the observation time of motion was larger because it did not correspond to the observer's expectations [20]. There are many studies on human motion, for example on human gait. In [28], subjects are able to recognize the emotional state (joy, sadness, etc.) from point-light stimuli. These studies show that biological motion convey higher-level information (contextual, emotional, etc.). The human brain uses low-level motion-detecting processes to recognize biological movements. It is assumed that it makes perception more pleasant and economical from a cognitive effort point of view [19].

It therefore seems essential to design motion models as 'biological' as possible in order to generate animations that are perceptually acceptable. For that, studies on motion should determine how these laws are expressed in SL. Previous work showed that there can be a 'SL effect' on these laws [2], [3]. We are still far from having enough knowledge to be satisfied with purely synthetic animations. Another possible approach is to replay movements that were recorded on people.

3 ANIMATING AVATARS USING MOTION REPLAY

Virtual signers can be animated using motion capture (mocap) on real actors [9], as it provides highly realistic and comprehensible content. Thanks to mocap, the 3D trajectories of main articulations of the body are recorded and then mapped to the skeleton of a virtual character. With such accurate systems, we solve the problem of perceptual acceptability.

However, as stated above, the motion may also convey rich information about a person. As an example, it can make the actor identifiable, such as the voice can allow a given person to be recognized. As previously stated, humans can extract important information from biological motion, such as one's intentions, emotions, or identity. How such information can be retrieved from complex movements remains a challenging question for both psychology and computer vision areas. Extraction of human traits from a complex signal is a fundamental problem that has been addressed in different domains. In the auditory field, studies investigated the perception of extra-linguistic cues in speech [1], [23] and notably the recognition of a particular speaker [14]. Similar approaches have been used in the visual domain to study the categorization and identification of human faces [27], [26], [22]. Studies of Johansson [11], [12] addressed the question for human motion, introducing the notion of point-light (PL) stimuli. This display separates information given by dynamic cues from characteristics such as shape or aspect of the person. Using this device, Johansson showed that humans could recognize a set of moving dots as a human walker. Point-light displays are widely used since then. Different studies demonstrated that they contain enough information to recognize familiar people from their gaits [6], [16], [10], [25].

As mocap systems capture equivalent information as point-light displays, such information might be perceived through animations made with motion replay. We have observed on several occasions that some LSF (French Sign Language) signers were able to identify the person who had been registered with mocap, in the case of long utterances performed by an humanoid avatar, such as in alert messages in French train stations [24], and even in the case of isolated signs such as those of the serious game SignEveil by MocapLab [21], with avatars such as bear or cat. The need for producing messages anonymously is an important demand of Deaf people for some applications, such as posting anonymous content on blogs or social networks. This entails studying avatar animations from a perception point of view. Some studies are analysing SL mocap data from a linguistic point of view [17], [18], [5] but as we know, no one studied the extra-linguistic cues conveyed by such signals. This is a question that we are beginning to study at LIMSI.

4 ONGOING STUDIES ON IDENTITY AND ANONYMIZATION

We have started addressing the question of gestural identity in LSF motion [4]. The aim of this work is to find critical features which differentiate the gestural identity of human signers in order to provide a better control of the animated signers. This multidisciplinary research yields contributions to psychology, computational science and LSF animation with both fundamental and applied perspectives. On the one hand, it helps us better understand how critical information can be extracted from the complex SL motion stimulus. On the other hand, it enables controlling social and human characteristics of the signing avatar regarding motion. For example, manipulating gestural identity would allow to produce signing animation 'in the style of' someone or to anonymize our own signing. Anonymizing the gesture of replayed motion can add substantial contribution to the existing virtual signing systems which only anonymize shape of the agent for now.

To undertake this work, we have adopted a pluridisciplinary approach, combining perception studies and computing methods. In a first step, we develop evaluation methods for the identification of signers from their gestures. The aim is to assess whether the identity of a signer can be transmitted through the movement only. To this aim, we present videos of real signers to the participants, shown as point-light displays. Then, we apply machine learning techniques for the analysis of mocap data [15] in a way to extract human features from LSF motion (e.g. identity, age, emotions...). Based on the same stimuli, we run a classifier using a similar approach to the one developed for face, voice and gait recognition.

This ongoing work also motivated the design of a new dedicated LSF mocap corpus that will be carried out in the coming months.

REFERENCES

[1] Belin, P., Bestelmeyer, P. E., Latinus, M., and Watson, R. Understanding voice perception. *British Journal of Psychology 102*, 4 (2011), 711–725.

[2] Benchiheub, M., Berret, B., and Braffort, A. Collecting and analysing a motion-capture corpus of french sign language. In *7th International Conference on Language Resources and Evaluation - Workshop on the Representation and Processing of Sign Languages (LREC-WRPSL 2016)*, ELRA, pp. 7–12.

[3] Benchiheub, M.-E.-F. *Contribution Ãă l'analyse des mouvements 3D de la Langue des Signes Franaise (LSF) en Action et en Perception.* PhD thesis, University of Paris-Saclay, Orsay France, 2017.

[4] Bigand, F., Prigent, E., and Braffort, A. Animating virtual signers: The issue of gestural anonymization. In *Proceedings of the 19th ACM International Conference on Intelligent Virtual Agents* (2019), ACM, pp. 252–255.

[5] Catteau, F., Blondel, M., Vincent, C., Guyot, P., and Boutet, D. Variation prosodique et traduction poétique (lsf/français): Que devient la prosodie lorsqu'elle change de canal? In *Journées d'Étude sur la Parole* (2016), vol. 1, pp. 750–758.

[6] Cutting, J. E., and Kozlowski, L. T. Recognizing friends by their walk: Gait perception without familiarity cues. *Bulletin of the psychonomic society 9*, 5 (1977), 353–356.

[7] Filhol, M., and Mcdonald, J. Extending the azee-paula shortcuts to enable natural proform synthesis.

[8] Filhol, M., McDonald, J., and Wolfe, R. Synthesizing sign language by connecting linguistically structured descriptions to a multi-track animation system. In *International Conference on Universal Access in Human-Computer Interaction* (2017), Springer, pp. 27–40.

[9] Gibet, S. Building french sign language motion capture corpora for signing avatars. In *Workshop on the Representation and Processing of Sign Languages: Involving the Language Community, LREC 2018* (2018).

[10] Jacobs, A., Pinto, J., and Shiffrar, M. Experience, context, and the visual perception of human movement. *Journal of Experimental Psychology: Human Perception and Performance 30*, 5 (2004), 822.

[11] Johansson, G. Visual perception of biological motion and a model for its analysis. *Perception & psychophysics 14*, 2 (1973), 201–211.

[12] Johansson, G. Spatio-temporal differentiation and integration in visual motion perception. *Psychological research 38*, 4 (1976), 379–393.

[13] Kipp, M., Heloir, A., and Nguyen, Q. Sign language avatars: Animation and comprehensibility. In *Intelligent Virtual Agents* (2011), Springer, pp. 113–126.

[14] Kuhn, R., Nguyen, P., Junqua, J.-C., Goldwasser, L., Niedzielski, N., Fincke, S., Field, K., and Contolini, M. Eigenvoices for speaker adaptation. In *Fifth International Conference on Spoken Language Processing* (1998).

[15] LIMSI, and CIAMS. Mocap1, 2017. https://hdl.handle.net/11403/mocap1/v1.

[16] Loula, F., Prasad, S., Harber, K., and Shiffrar, M. Recognizing people from their movement. *Journal of Experimental Psychology: Human Perception and Performance 31*, 1 (2005), 210.

[17] Malaia, E., Borneman, J., and Wilbur, R. B. Analysis of asl motion capture data towards identification of verb type. In *Proceedings of the 2008 conference on semantics in text processing* (2008), Association for Computational Linguistics, pp. 155–164.

[18] Malaia, E., Wilbur, R. B., and Milković, M. Kinematic parameters of signed verbs. *Journal of Speech, Language, and Hearing Research 56*, 5 (2013), 1677–1688.

[19] Mather, G., Radford, K., and West, S. Low-level visual processing of biological motion. *processing of biological motion 24*, 1325 (1992).

[20] Méary, D., Kitromilides, E., Mazens, K., Graff, C., and Gentaz, E. Four-day-old human neonates look longer at non-biological motions of a single point-of-light. *PLoS ONE 2(1): e186.*

[21] MocapLab. Signéveil. http://www.mocaplab.com/fr/projects/signeveil/.

[22] O'Toole, A. J., Abdi, H., Deffenbacher, K. A., and Valentin, D. Low-dimensional representation of faces in higher dimensions of the face space. *JOSA A 10*, 3 (1993), 405–411.

[23] Schweinberger, S. R., Kawahara, H., Simpson, A. P., Skuk, V. G., and Zäske, R. Speaker perception. *Wiley Interdisciplinary Reviews: Cognitive Science 5*, 1 (2014), 15–25.

[24] SNCF, Websourd, and LIMSI. Jade. https://www.accessibilite.sncf.com/la-demarche-d-accessibilite/equipements/amenagements-en-gare/s-informer/article/jade.

[25] Troje, N. F., Westhoff, C., and Lavrov, M. Person identification from biological motion: Effects of structural and kinematic cues. *Perception & Psychophysics 67*, 4 (2005), 667–675.

[26] Turk, M. A., and Pentland, A. P. Face recognition using eigenfaces. In *Proceedings. 1991 IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (1991), IEEE, pp. 586–591.

[27] Valentin, D., Abdi, H., and O'Toole, A. J. Categorization and identification of human face images by neural networks: A review of the linear autoassociative and principal component approaches. *Journal of biological systems 2*, 03 (1994), 413–429.

[28] Venture, G., Kadone, H., Zhang, T., GrÃĺzes, J., Berthoz, A., and Hicheur, H. Recognizing emotions conveyed by human gait. *International Journal of Social Robotics volume=6, number=4, pages=621–632* (2014).