Perspective-based Character Description in Interactive 3D Environments

Gonzalo Méndez[1], **Raquel Hervás**[2], **Pablo Gervás**[1], **Ricardo de la Rosa**[2], **Daniel Ruiz**[2]

[1] *Instituto de Tecnología del Conocimiento - Universidad Complutense de Madrid*
{gmendez,pgervas}@ucm.es

[2] *Facultad de Informática - Universidad Complutense de Madrid*
raquelhb@fdi.ucm.es,{ricarosa,danruiz}@ucm.es

Interactive narrative generation in 3D environments can benefit from the fact that a lot of information is implicitly and explicitly available in the environment to generate perspective based descriptions of people and objects that change according to the user's point of view, a problem that has not been addressed in general in the field of referring expression generation [5]. One of the difficulties when generating these descriptions is deciding when we must change the referring expression we use to describe the element's situation, even more if we take into account that not everybody refers to an element in the same manner.

A survey was carried out to investigate how spatial properties and points of view affect the composition of descriptions used when referring to another person. With the results of this survey, and using the graphical engine Unity 3D, an algorithm has been developed to generate descriptions of characters in real time that change according to the user's position within the environment. Further details about this work can be read in [7].

The developed algorithm was implemented in a game where the user had to find the character that was being described, for which he could move around the environment and the provided description changed accordingly. The content of the description is based on the character's physical appearance, its position within the environment, and its situation with respect to other relevant characters and objects that are present in the environment: *(1) attribute-based description*, which refers to the static characteristics of an individual and its environment, and they cannot change during the simulation. This part of the description is generated using classical referring expression generation algorithms, such as *Greedy* or *Incremental* [5]; *(2) perspective-based description*, which has to be generated in real time according to the situation of the user relative to the situation of the described person.

Each description can be composed of sub-descriptions, which contain partial information of a full description. Every sub-description is generated according to the data that is obtained from the scene. There are four possible types of sub-descriptions: *(1) meta description*: the static description generated by the classical algorithms; *(2) description of the static points*: this description contains the information of reference points scattered all over the scene; *(3) description of the visibility*: it contains the information about the visibility of the described person from the user's point of view. For example, the fact that the described person is behind a column; *(4) distance between the described person and the user*: it contains a textual explanation of the distance between the described person and the user.

This paper has been partially supported by the project ConCreTe (611733), funded by the European Commission under FP7, the ICT theme, and the Future Emerging Technologies (FET) program.

The last three sub-descriptions are generated and stored until the algorithm checks certain flags and the information needed for the description is generated accordingly. First it checks the distance between the user and the reference points previously put into the scene. If the distance to the user is greater than a predefined constant, the generated description contains information about the proximity of the described person to that point. Otherwise, the algorithm checks if the described user is on the camera area. If not, the generated description must contain the positional references of the described person to the user: it indicates if the described person is to the left, right or behind the user. If the described person is within the camera area, it checks the absolute distance between the described person and the user. It indicates the user whether he is near, not far or far from the described person. By mixing sub-descriptions and flag checking, the perspective-based descriptions are generated. As a result, in each case a specific order and sub-description appear in the generated description (e.g. *The boy in the black shirt sitting near the window. The described person is behind another person. He is not far*).

The proposed solution to generate descriptions is not currently based on the use of fuzzy logic, although it can benefit from it to generate descriptions of spatial relationships. Although we have not been able to find an approach of this kind in the reviewed literature, there have been some efforts to solve similar problems in the fields of image analysis and computer vision, as described in [2], whose authors have subsequently developed several methods to describe spatial relations between objects [1, 3]. Other authors have tackled the problem of automatic scene descriptions in image analysis using fuzzy rule-based systems [4] and fuzzy sets [6], through the use of histograms of angles and forces and a dictionary of labels.

References

[1] I. Bloch. Fuzzy relative position between objects in image processing: new definition and properties based on a morphological approach. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 7(2):99 – 133, April 1999.

[2] I. Bloch and A. Ralescu. Directional relative position between objects in image processing: a comparison between fuzzy approaches. *Pattern Recognition*, 36(7), 2003.

[3] C. Hudelot, J. Atif, and I. Bloch. Fuzzy spatial relation ontology for image interpretation. *Fuzzy Sets and Systems*, 159(15):19291951, August 2008.

[4] J. M. Keller and X. Wang. A fuzzy rule-based approach to scene description involving spatial relationships. *Comp. Vision and Image Understanding*, 80(1):21–41, Oct. 2000.

[5] E. Krahmer and K. van Deemter. Computational generation of referring expressions: A survey. *Computational Linguistics*, 38(1):173–218, March 2012.

[6] P. Matsakis, J.M. Keller, L. Wendling, J. Marjamaa, and O. Sjahputera. Linguistic description of relative positions in images. *IEEE Transactions on Systems, Man and Cybernetics*, 31(4):573–88, 2001.

[7] R. de la Rosa and D. Ruiz. Perspective-based referring expression generation in a 3D environment. BSc. Project, Facultad de Informática - UCM, June 2015.