USING FUZZY LOGIC TO MANAGE UNCERTAIN MULTI-MODAL DATA IN AN ARCHAEOLOGICAL GIS

C. De Runz[a,b,*], **E. Desjardin**[a], **F. Piantoni**[b], **M. Herbin**[a]

[a] CReSTIC EA-3804, IUT Reims Chalons Charleville, rue des Crayères, BP 1035, 51687 Reims cedex 2, France
(cyril.de-runz, eric.desjardin, michel.herbin)@univ-reims.fr
[b] HABITER EA-2076, URCA, 57 rue Taittinger, 51096 Reims cedex, France
(cyril.de-runz, frederic.piantoni)@univ-reims.fr

KEY WORDS: Fuzzy Logic, Archaeological Data, Fuzzy Multi-Modal Data, Fuzzy Hough Transform, Uncertain Merging

ABSTRACT:

Into an archaeological GIS, stored information is in spatio-temporal multi-modal form. Dates are estimated according to excavation data. Localizations are determined by several tools such as GPS, radar,... All of these imply some uncertainty, for which fuzzy logic becomes a common use. By representing data by fuzzy number depending on specific modalities, we obtain fuzzy spatio-temporal data. In this paper, we propose to illustrate the interest of fuzzy data representation in archaeology.
As Roman streets in Reims usually form segments of lines, we apply the fuzzy Hough transform to those fuzzy multi-modal data to estimate their presence in accordance with given periods. According to experts, we propose to visualize in GIS archaeological data and simulations obtained by merging this fuzzy information.

1 INTRODUCTION

According to Burrough and McDonnell (Burrough and McDonnell, 1998), the management of data quality into a GIS is fundamental for data exploitation and the trust on results we have. Geographic Information Systems (GIS) usually store multi-modal data and employ them in spatial and image analysis. Archaeological data often stem from excavations and group date, localization, etc. Localizations are determined by several tools. Each tools imply some imprecision. Furthermore, the past localization, and the actual localization could be different due to soil movements. Dates generally spring from expert interpretations or estimations, depending on the excavation context. As numerical representations of linguistic period codification vary in accordance with experts, data are subject to part of uncertainty. Moreover, except for localizations, data are lacunar.

The Fuzzy Sets theory, introduced by Zadeh (Zadeh, 1965), proposes a formalism to represent uncertain knowledge. A review of methods to manage data quality in geographic information are proposed in (Goodshild and Jeansoulin, 1997). The use of Fuzzy Logic tends to become classical in GIS (Altman, 1994). In archaeology, the modeling by fuzzy sets of uncertain data is not classical. In this paper, excavations data are fuzzified to take care about their uncertain aspect. So stored multi-modal data become after the fuzzification process fuzzy multi-modal data. We illustrate the interest in archaeology of the fuzzy data representation by a simulation.

Using the fuzzy logic and the fuzzy Hough transform (FHT), we propose a simulation process of Roman streets prediction into the archaeological SIGRem project (Pargny and Piantoni, 2005), an archaeological GIS dedicated to Roman periods of Reims, using ”BDRues” geodatabase. Data contain three kinds of information : the localization, the orientation, and the date. FHT is a classical method for detecting straight lines in images. Because Roman streets in Reims form segment of lines, we propose to use it on “BDRues”. We apply FHT on each fuzzy feature of excavation data, and we merge the information contained into the three FHT accumulators into a new one. The application gives an estimation of the potential presence of Roman streets for a given period. Using this fusion function, we visualize the results into a GIS software.

In this paper, we present the archaeological data and their fuzzy representation. The management of archaeological data using fuzzy logic and possibility theory within archaeological environment is explained. We propose a simulation process on fuzzy multi-modal data using a merging function. Finally, the results of the simulation are exposed.

2 ARCHAEOLOGICAL DATA

2.1 Multi-modal and uncertain data

Conolly and Lake (Conolly and Lake, 2006) emphasize the importance of storing the error associated with data. In fact, our excavation data contain at least the localization and usually the date and the information about the orientation.

The first one depends on tools and methodologies used by archaeologist during excavations. For example, data localizations acquired with ancient tools during the 18^{th} Century are less precise than nowadays with GPS. Likewise the GPS localization is subject to lack of precision. Even if the localization is perfect, the coordinates represent the present localization. And year after year, objects could have been shift by soil movements or local transport, which could only be estimated. Thus the localizations present part of uncertainty.
Dates generally stem from experts interpretations or estimations, depending on the excavation context. Furthermore, numerical representations of linguistic periods codification vary in accordance with experts.
When we have an information about orientation, this one is generally issued from expert analysis on materials found during excavations. Unfortunately, there is no record of error range.

Even if all those features are important for us some of them are sometimes missing. Thus we also have to deal with lacunar data.

Many papers as (Devillers and Jeansoulin, 2006) deal with the quality of information in GIS. The literature discusses imprecise reasoning (Guesgen and Albrecht, 2000). Altman (Altman, 1994) proposes to use Fuzzy Set theoretic approaches for handling imprecision in spatial analysis. This approach is applied to understand phenomena as road traffic, dynamic processes, or contamination of soil (Dixon, 2005, Dou et al., 1999, Dragicevic and Marceau, 2000, Mitra et al., 1998). The use of fuzzy logic becomes classical to manage uncertain data in simulation processes (Goodshild and Jeansoulin, 1997) in GIS.

As explain before, the multiplicity of sources and the confidence in data imply that archaeological data are multi-modal and uncertain. We propose to manage those data using fuzzification processes. For example, a fuzzy localization is associated to the actual localizations of excavations extracted from our geodatabase (similar case for date and orientation).

2.2 Fuzzy representation

In this work, some models of representation by fuzzy numbers are used illustrated by Figures 1, 2 and 3. In Figure 1, we propose to associate a fuzzy localization ($fLoc_{ep}$) to an excavation point ep. Near the excavation point coordinates, the fuzzy membership function is equal to one and decreases when distance to ep grows up. Figure 2 presents a model for the fuzzy orientation ($fOrien_{ep}$), where the fuzzy membership function decreases when the angle difference with the stored orientation grows up. Figure 3 illustrates the fuzzy period $[bd, ed]$ ($fDate_{ep}$).

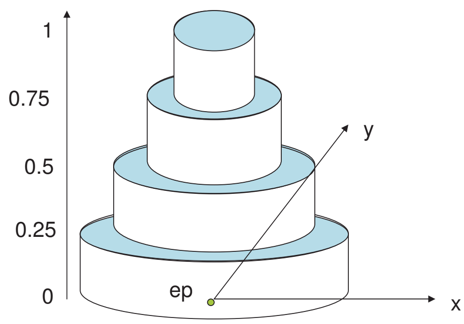

Figure 1: Fuzzy membership function $fLoc_{ep}$

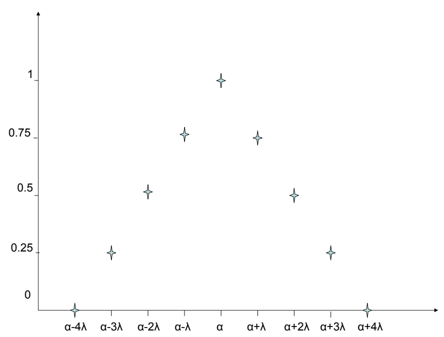

Figure 2: Fuzzy membership function $fOrien_{ep}$

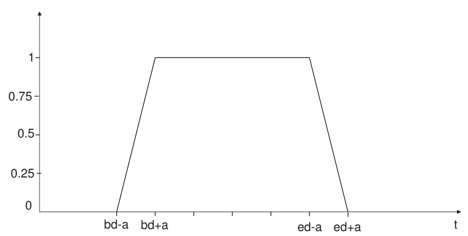

Figure 3: Fuzzy membership function $fDate_{ep}$

This association of fuzzy numbers to each feature of our excavation data allows us to manage them with their uncertainty into simulation process.

During Roman age, cities where built in accordance with a regular grid arrangement of streets. So Roman streets had the particularity to be close to lines. In the next section, we propose to use the fuzzy Hough transform (FHT) (Han et al., 1994) concept to simulate city maps according to Roman periods.

3 SIMULATION INTO A ARCHAEOLOGICAL GIS

3.1 Presentation of the simulation

One of the most powerful methods for the detection and recognition of patterns with known simple geometrical shapes (as lines) in images is the Hough transform (HT) (Duda and Hart, 1972, Hough, 1962, Illingworth and Kittler, 1988). HT could be generalized and unsupervised for detecting natural shapes (Bonnet, 2002). To take care of noise and quantization error, the binary aspect of the decision making can be smoothed as proposed by Han et al. (Han et al., 1994) using their extension of HT to the fuzzy logic (fuzzy Hough transform - FHT). For each candidate in the image, they define a fuzzy neighborhood within which each point contributes to the potentiality of line presence.

Our application has to work with fuzzy multi-modal data. Data contain three features (localization, orientation and date). Roman streets usually form segment of lines. We propose to apply FHT to the geodatabase. This constraint implies to use three FHT: one for the localization, one for the orientation and the last one for the correspondence of dates with the given period. As only one valuation is wanted by experts, we need to fusion the three FHT results to build the final estimation. The application estimates Roman streets for a given period and permits us to visualize them into a GIS.

3.2 Principle of the fuzzy Hough transform

We briefly introduce the principle of Hough transform (Hough, 1962, Duda and Hart, 1972). Surveys of HT can be found in (Illingworth and Kittler, 1988, Leavers, 1993). HT was originally defined to detect straight lines in binary images. It is a mapping from image space to parameter space which is represented by an accumulator array. In the case of straight lines, the accumulator array dimensions are (ρ, θ) where θ is the angle of the straight line and ρ its distance to the origin (see in Fig 4).

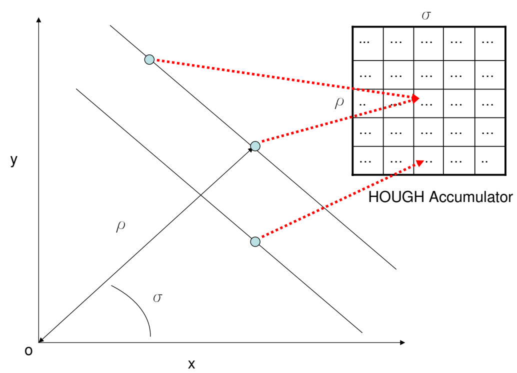

Figure 4: Hough transform principle

Each point of interest votes, incrementing the accumulator array cell corresponding to each line potentially going through it. If lines are present in the image we can detect them by finding cells with greatest scores.

Han et al. (Han et al., 1994) propose with the Fuzzy Hough transform to detect lines in images with noise, or quantization error. FHT takes into consideration points and their neighborhood. Han et al. use a fuzzy set to build a neighborhood matrix of the studied point. This matrix determines the value of the increment of the counter function $FHTAcc$.

In our application, FHT is not used with images but with multi-modal data stored in a geodatabase. In the next section, a method to determinate Roman streets according to a given period is proposed, and this method uses FHT with fuzzy multi-modal data.

3.3 Three fuzzy Hough transforms accumulators to estimate Roman Streets according to a period

The goal of the simulation is to estimate the potentiality of streets presence in antique Reims (Durocortorum) according to a specific period. Roman streets have the particularity to be most of time linear. The idea is to apply fuzzy Hough transform to each feature using their associated fuzzy numbers.

Excavation points are the points of interest. Data are characterized by three features, thus three FHT are needed. The first one ($FHTLoc$) is devoted to the localization, the second one to the orientation ($FHTOrien$) and the last one to the correspondance with a given date ($FHTDate$).
For an excavation point *ep*, in presence of all features, each point $p \in support(fLoc_{ep})$ votes for all lines (ρ, θ) going through it with $\theta \in support(fOrien_{ep})$.
The value of the vote in the accumulator $FHTLocAcc$ of $FHTLoc$ is $fLoc_{ep}(p)$. For $FHTOrien$, the increment in $FHTOrienAcc$ is equal to $fOrien_{ep}(\theta)$. To obtain the street for the given period gp, the counter function $FHDateAcc$ of $FHTDate$ is increased by $max(min(gp, fDate_{ep}))$.
When we do not have the orientation, then we vote for all lines going through p in $FHTLoc$ and $FHTDate$. If the information about date is missing, then only $FHTLocAcc$ and $FHTOrienAcc$ are incremented. Next the vote processes, each counter cell is normalized by the maximum of the associated counter function results.

To generate maps based on fuzzy multi-modal data, we need to merge those three FHT. The next section deals with the choice of the merging function.

3.4 A fusion function to visualize simulation into a GIS

In this section the three FHT ($FHTLoc$, $FHTOrien$ and $FHTDate$) are considered as fuzzy sets. The associated membership functions are the FHT accumulators ($FHTLocAcc$, $FHTOrienAcc$ and $FHTDateAcc$).

In fusion of information context, merging uncertain data is a classical problem. Dubois and Prade (Dubois and Prade, 2004) remark that the choice of a fusion mode depends on the nature of the items to merge and the representation framework.

A review of traditional aggregation and fusion operators can be found in (Detyniecki, 2000). Classical operators as *t-norm*, *t-conorm* in fuzzy context (for Zadeh the Min and Max) consider the order of data has no importance (symmetric function). Furthermore, $t - norm$ and $t - conorm$ admit a neutral value and an absorbent element.

In our context, each feature is of importance but does not have the same influence. In fact, $FHTDateAcc$ cell values could be more important than $FHTLocAcc$.

Thus we need a merging function with no neutral value and no annihilator, and the function is not symmetric. In classical merging, by extending the arithmetic mean, the weighted mean can solve this problem. In fuzzy context, the fusion function corresponding to the weighted mean is a parameterized $Choquet$ function (Choquet, 1953). The merging function is defined as follows:

$$\begin{aligned} final = \quad & \lambda * FHTLocAcc \\ & +\mu * FHTOrienAcc \\ & +\nu * FHTDateAcc, \end{aligned}$$

where the weights λ, ν and μ are non negative and $\lambda + \nu + \mu = 1$.

The weights (λ, μ, ν) value depends on applications and/or experts goals. Using this function, the three FHT could be reduced to one fuzzy set $Final$ (the associated membership is $final$). Thus, we apply an α-cut to select the lines (potential presence of streets). We use an other function to reduce line to fuzzy segment.

This simulation process is used to estimate the potential of Roman streets presence for periods defined by users. The next section presents results obtained by the method, which are compared with maps delivered by experts for the same period.

4 RESULTS

Into the SIGRem project (Pargny and Piantoni, 2005), we use the method to estimate the potential presence of streets at the Third Century BC in Reims.

The weights (λ, μ, ν) of $final$ are empirically evaluated as (3/13, 1/13, 9/13) in the application. This affectation permits to give three times more importance to date correspondence compared with localization, and three times more importance to localization in comparison to orientation.

The comparison between simulated maps (as Figure 5) and maps from experts (as Figure 6), validates the method. In these figures, simulated streets and streets defined by experts are most of the time similar.

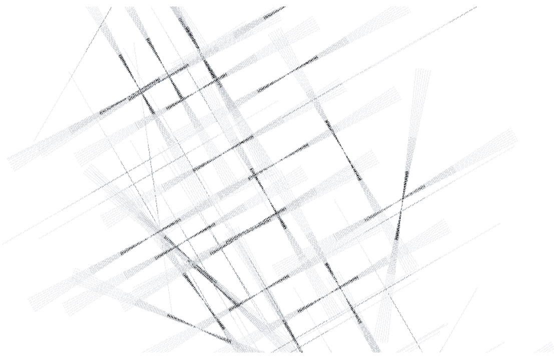

Figure 5: Simulated map of Reims during the third century

5 CONCLUSIONS AND FUTURE WORK

Archaeological GIS store excavation data. The localizations or orientations depend on tools, soil movements, transport, etc. The dates are only estimation given by experts. Thus excavation data include a part of uncertainty.

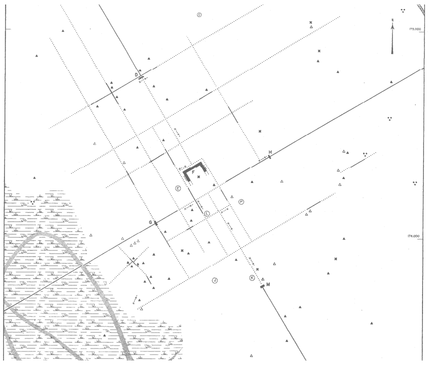

Figure 6: Map defined by experts of Reims during the third century

Management of uncertainty is fundamental to simulation processes in archaeology. Because archaeological data are uncertain on many way, fuzzy logic help us to manage them with their uncertainty. We apply fuzzy set representation to multi-modal archaeological data in a simulation process. This simulation aims to estimate the potentiality of a street presence.

Our simulation for Roman streets presence estimation, builds three FHT for localization, orientation and correspondence to a given date and aggregates multi-modal information by fusion. Our contribution improves the Fuzzy Hough Transform taking into account both the multi-modality and the inaccuracy and uncertainty of data. Comparison between simulated maps and maps from experts validates our method.

In our future work, by using our street presence estimator and information stored in GIS, we will study the fuzzy matching process between digitized image of ancient maps and archaeological data issued from excavations.

ACKNOWLEDGEMENTS

The authors would like to thank SRA (Regional Service in Archaeology) Champagne-Ardenne and INRAP (National Institute of Research in Preventive Archaeology) in Reims for the access to their data, and Dominique Pargny for his contribution to the SIGRem project.

REFERENCES

Altman, D., 1994. Fuzzy set theoretic approaches for handling imprecision in spatial analysis. Int. J. Geographical Information Systems 8 3, pp. 271–290.

Bonnet, N., 2002. An unsupervised generalized hough transform for natural shapes. Pattern Recognition 35, pp. 1192–1196.

Burrough, P. and McDonnell, R., 1998. principle of Geographical Information Systems. Oxford University Press.

Choquet, G., 1953. Theory of capacities. Annales de l'Institut Fourier 5, pp. 131–295.

Conolly, J. and Lake, M., 2006. Geographic Information System in Archaeology. Cambridge University Press.

Detyniecki, M., 2000. Mathematical aggregation operators and their application tovideo querying.

Devillers, R. and Jeansoulin, R. (eds), 2006. Data Space Quality: An Introduction. Iste Publishing Company.

Dixon, B., 2005. Groundwater vulnerability mapping: a gis and fuzzy rule based integrated tool. Applied Geography 20, pp. 1–21.

Dou, C., Wolt, W. and Bogardi, I., 1999. Fuzzy rule-based approach to describe solute transport in the unsaturated zone. Journal of Hydrology 220 1-2, pp. 74–85.

Dragicevic, S. and Marceau, D., 2000. An application of fuzzy logic reasoning for gis temporal modeling of dynamic processes. Fuzzy Sets and Systems 113, pp. 69–80.

Dubois, D. and Prade, H., 2004. On the use of aggregation operations in information fusion processes. Fuzzy Sets and Systems 142, pp. 143–161.

Duda, R. and Hart, P., 1972. Use of the hough transform to detect lines and curves in pictures. Comm. ACM 15 (1), pp. pp 11–15.

Goodshild, M. and Jeansoulin, R. (eds), 1997. Data Quality in Geographic Information. Hermes.

Guesgen, H. and Albrecht, J., 2000. Imprecise reasonning in geographic information systems. Fuzzy Sets and Systems 113, pp. 121–131.

Han, J., Koczy, L. and Poston, T., 1994. Fuzzy hough transform. Pattern Recognition Letters 15, pp. 649–648.

Hough, P., 1962. Method and means for recognizing complex patterns. US 3 069 654.

Illingworth, J. and Kittler, J., 1988. A survey of the hough transform. Information Control 44, pp. 87–116.

Leavers, V., 1993. Which hough transform. CVGIP 58, pp. 250–264.

Mitra, B., Scott, H. and McKimmey, J., 1998. Application of fuzzy logic to the prediction of soil erosion in a large watersheld. Geoderma 86, pp. 183–209.

Pargny, D. and Piantoni, F., 2005. Méthodologie pour la gestion, la représentation et la modélisation des données archéologiques. In: conférence Francophone ESRI.

Zadeh, L., 1965. Fuzzy sets. Information Control 8, pp. 338–353.