A COMPARISON OF NEURO-FUZZY AND TRADITIONAL IMAGE SEGMENTATION METHODS FOR AUTOMATED DETECTION OF BUILDINGS IN AERIAL PHOTOS

Thomas Knudsen[a] **Hamed Hamid Muhammed**[b] **Brian Pilemann Olsen**[a]

[a]Kort- & Matrikelstyrelsen, 8 Rentemestervej, DK-2400 Copenhagen NV, Denmark

mailto:{thk|bpo}@kms.dk http://research.kms.dk/∼{thk|bpo}

[b]Centre for Image Analysis, Uppsala University, Lagerhyddvägen 17, SE-752 37 Uppsala, Sweden

mailto:hamed@cb.uu.se http://www.cb.uu.se/∼hamed

KEY WORDS: neural, segmentation, classification, urban, photogrammetry

ABSTRACT

Using a set of colour-infrared aerial photos, we compare a newly developed neural net based clustering method with a method based on the classical ISODATA algorithm. The primary focus is on the detection of buildings and it shows that while the traditional method has an advantage in splitting background from foreground, the neural net based method results in a much more uniform segmentation of the input images.

1 INTRODUCTION

Spectral detection of buildings in aerial photos is a non-trivial task as the overall spectral characteristics of buildings are usually very varied and not easily distiguished from the background. A low level clustering step tends to improve the separability, and can be carried out using a multitude of different algorithms, varying from the classic ISODATA (Ball and Hall, 1965), to newer approaches as SUSAN (Smith and Brady, 1997), or the recent methods SYNERACT (Huang, 2002) and EDISON (Christoudias et al., 2002).

In this paper, we compare FC-WINN (Hamid Muhammed, 2002), a newly developed neural net based unsupervised clustering method with PRECLUST, another unsupervised method based on ISODATA. PRECLUST has previously shown useful as a preprocessor for colour-infrared (CIR) aerial photos in a building detection system (Knudsen and Olsen, 2002, Olsen et al., 2002)

2 PRECLUST

PRECLUST is based on ISODATA, the classic self organizing iterative clustering algorithm. ISODATA is simple to implement and readily available in most image analysis packages for remote sensing data. Its strengths (e.g. its simplicity) and weaknesses (e.g. its potential lack of convergence) are well known, and has been well documented in the literature, most recently in Huang's rationale for the SYNERACT method (Huang, 2002).

Hence, the method is not described in detail here, only a few remarks on the implementation used for PRECLUST is given. PRECLUST is intended as a preprocessor for aerial photos, delivering the training set for a following supervised classification stage. As we are mostly interested in detecting buildings, the clustering is carried out under a mask constructed from map database registrations of existing buildings (Knudsen and Olsen, 2002). The entire image is afterwards segmented using the clusters from the iteration process, but a threshold is applied such that pixels situated spectrally far from the clusters generated are assigned to a background class.

To speed things up, we simplify the process compared to the original ISODATA method by using minimum distance, rather than maximum likelihood as the clustering parameter. But compared to the most basic ISODATA implementation, where the number of classes is pre-selected and fixed, we complicate matters a bit by allowing clusters to split and/or combine for each iteration. This may increase the risk of no convergence, but also increases the chance of getting a set of resulting cluster centres that fit with the dataset at hand.

3 FC-WINN

FC-WINN is a neuro-fuzzy system based on a new type of artificial neural networks, called Weighted Incremental Neural Networks (WINN), which are introduced and discussed by Hamid Muhammed (2002) (see also section 3.1, below).

The new clustering algorithm follows a three-steps approach, where the input data set is first processed by the WINN to get the corresponding weighted connected net, which reflects and preserves the topology of the input data set, while the dimensionality of the problem is reduced considerably.

The second step is to cluster the resulting weighted connected net, using a watershed-like procedure (cf. e.g. Vincent and Soille (1991) for more details about watersheds). This simplifies the problem even more and makes it one-dimensional. The result of this procedure is a number of separated weighted connected sub-nets representing the obtained clusters, one sub-net for each cluster, where all nodes in a sub-net have the same label value.

Finally, in the third step, the clustering result is mapped onto the input data set, using a nearest neighbour classifier (Dasarathy, 1991), and each input data sample is classified as belonging to the nearest sub-net; i.e. the nearest cluster.

This approach has the great benefit that clustering the resulting weighted connected net instead of the input data set

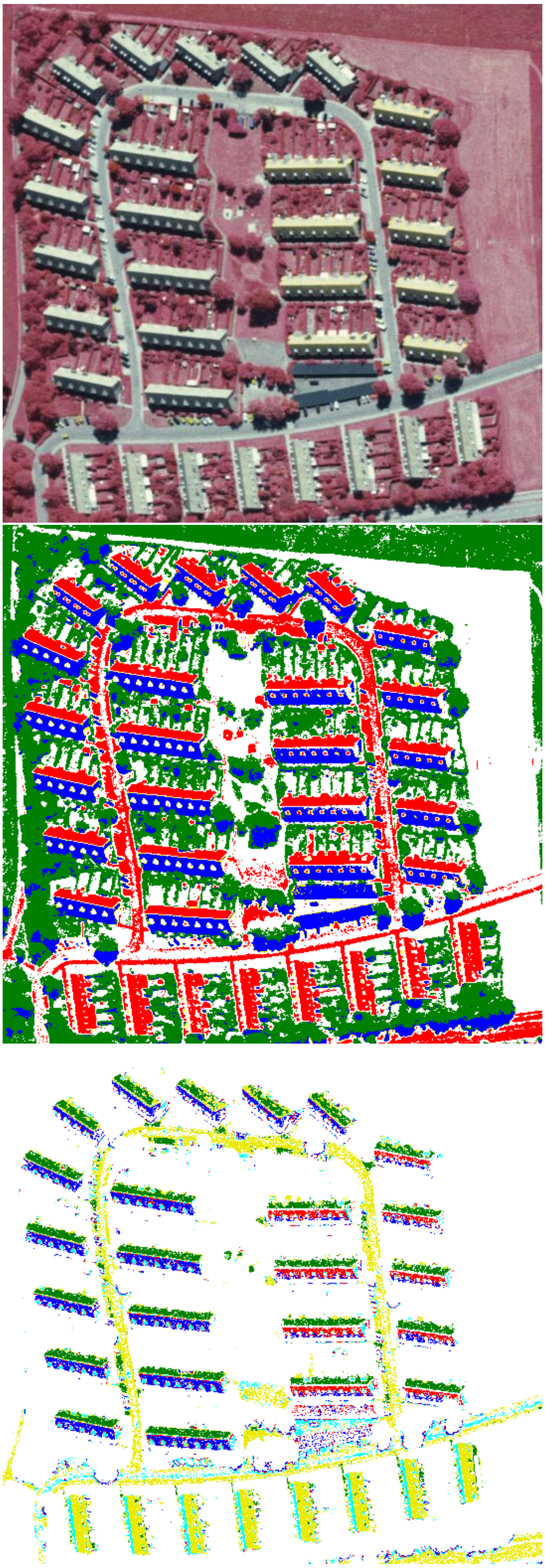

Figure 1: *Test area 1,* TOP: *Original colour-infrared image,* CENTER: *Segmentation using FC-WINN,* BOTTOM: *Segmentation using PRECLUST,*

Figure 2: *Test area 2,* TOP: *Original colour-infrared image,* CENTER: *Segmentation using FC-WINN,* BOTTOM: *Segmentation using PRECLUST,*

itself, makes it possible to reduce the memory and computational load considerably in the case of large input data sets, such as gigabyte sized full-resolution aerial photos. The obvious reason is the limited number of nodes in the resulting weighted connected net.

3.1 WINN

WINN (Hamid Muhammed, 2002), is an incremental self-organising model (Fritzke, 1996) with no pre-defined structure, and therefore no restrictions on the dimensionality of the input data set, which can have different dimensionalities in different regions of the input space. The model is built-up by successive addition, adaptation, and sometimes deletion of elements (i.e. nodes and edges), according to suitable strategies, until a stopping criterion is met.

A weighted connected net, consisting of weighted nodes connected by weighted edges, is produced, where the weights are proportional to the local densities of data samples in input space. The resulting net preserves the topology of the input data set.

The basic idea of the WINN algorithm is to generate and distribute a number of weighted nodes connected by weighted edges in the input data space, so that a relatively high weight-value corresponds to a relatively high density of input data samples in a neighbourhood around the corresponding node or edge, and vice versa. The algorithm begins with only two nodes connected by an edge, then new nodes and edges are generated and the old ones are updated (and sometimes deleted) while the learning process proceeds until a certain stopping criterion is met.

A fuzziness factor is introduced in the resulting weighted connected net, by propagating the influence of the input signal (which is the input data sample, currently presented to the neural network) to the *n* nearest nodes in the net, by updating them according to the signal-value, and by establishing and updating edges between the first winner, which is the first nearest node to the signal, and the other *n-1* nearest nodes to the signal.

The higher *n*-value is chosen, the higher connectedness of the resulting net is obtained, and consequently, the fuzzier the system becomes. The reason is the reduction of both of the between clusters distances and the within cluster distances, so that the clusters tend to merging into larger ones. On the other hand, at a certain *n*-value, the fuzziness of the system can be determined by choosing the number of the nodes in the resulting weighted connected net. Using fewer nodes in the net seems to correspond to increasing the fuzziness of the system. In other words, choosing a denser net (i.e. with more nodes) produces more and, consequently, relatively smaller clusters.

4 RESULTS & DISCUSSION

The results from segmentation of two test areas are shown in figures 1 and 2. A rough description of the contents of each of the resulting classes is given in table 1 (as both

```
-------------------------------------------

AREA1, FC-WINN:
    4 classes, total, n=2, nodes=500
    1 background/low vegetation     (white)
    1 trees/tall vegetation         (green)
    1 roads/illum. pt of bldgs        (red)
    1 shady part of buildings        (blue)
-------------------------------------------

AREA1, PRECLUST:
    6 classes, total
    1 background/vegetation         (white)
    2 roads/buildings         (yellow,cyan)
    1 illuminated part of buildings (green)
    2 shady part of buildings    (blue,red)
-------------------------------------------

AREA2, FC-WINN:
    6 classes, total, n=2, nodes=400
    1 background/low vegetation     (white)
    1 trees/tall vegetation         (green)
    1 roads                           (red)
    3 buildings          (blue,cyan,brown)
-------------------------------------------

AREA2, PRECLUST:
    9 classes, total
    1 background/vegetation         (white)
    1 roads                         (black)
    1 roads/illum. pt of bldgs     (yellow)
    1 roads/shady  pt of bldgs       (blue)
    1 shady part of buildings        (cyan)
    4 buildings      (green,pink,red,grey)
-------------------------------------------
```

Table 1: *A rough overview of the contents of the classes resulting from the unsupervised clusterings*

methods are unsupervised, the actual contents of the resulting classes is unknown until the conclusion of the segmentation).

The most striking difference between the two methods is that FC-WINN manages to reduce the number of classes to two-thirds of what PRECLUST needs. This is even more striking as PRECLUST takes the advantage of clustering registered building pixels only, and assigning input pixels far away from the resulting class centres to the background class.

The background class trick used by PRECLUST obviously makes it more stable than FC-WINN in discerning man made surfaces from vegetation covered surfaces (which is one of the major strengths in working with a near-infrared channel). It is, however, interesting that although only pixels registered as buildings were used in the clustering process, also roads get mixed up as building-like structures in the final classification. The mix-up appears for both PRECLUST and FC-WINN.

In general, FC-WINN creates much more uniform segments (virtually all of the road network in test area 2 ends in one segment, making it very easy to discard in a following post-classification step). PRECLUST segments are more noisy (partially due to the larger number of classes generated)—an extreme case is the two covered parking lot buildings in the lower right of test area 1: here PRECLUST generates a very noisy combination of segments belonging to a number of different classes, while FC-WINN gets both of the buildings into one class and one segment.

The noisy combination of small segments is, however, not a problem in the original application (i.e. building detection) of PRECLUST (Knudsen and Olsen, 2002), where the distinction between foreground and background is more important. As FC-WINN requires considerably more processor time than PRECLUST (at least in FC-WINN's current incarnation, where it is not optimized for speed), we still prefer PRECLUST for the purpose it was designed for. As a more general clustering package, however, FC-WINN seems stronger. If needed, the background trick can be fitted onto FC-WINN. In high-dimensional cases, where the phase space becomes more sparsely populated, FC-WINN may have an even higher advantage.

5 CONCLUSION

We have compared two clustering methods—one based on a traditional self organizing method, and one based on a combined neuro-fuzzy approach. The traditional method was most succesful in discerning foreground from background, while the neuro-fuzzy method was more succesfull in all other aspects. The relative merits normalized by required processor power is as yet not relevant, as the implementation of the neuro-fuzzy algorithm is still highly experimental and not optimized for speed.

REFERENCES

Ball, G. H. and Hall, D. J., 1965. A novel method of data analysis and pattern classification. Technical report, Stanford Research Insitute, Menlo Park, CA.

Christoudias, C. M., Georgescu, B. and Meer, P., 2002. Synergism in low level vision. In: 16th International Conference on Pattern Recognition. Track 1: Computer Vision and Robotics, Quebec City, Canada, p. 6.

Dasarathy, B., 1991. Nearest Neighbor (NN) Norms: NN Pattern Classification Techniques. IEEE Computer Society Press.

Fritzke, B., 1996. Growing self-organizing networks—why? In: ESANN'96: European Symposium on Artificial Neural Networks, pp. 61–72.

Hamid Muhammed, H., 2002. A new unsupervised fuzzy clustering algorithm (FC-WINN) using the new weighted incremental neural network (WINN).

Huang, K. Y., 2002. A synergistic automatic clustering technique (SYNERACT) for multispectral image analysis. Photogrammetric Engineering & Remote Sensing 68(1), pp. 33–40.

Knudsen, T. and Olsen, B. P., 2002. Detection of buildings in colour and colour-infrared aerial photos for semi-automated revision of topographic databases. This Volume.

Olsen, B. P., Knudsen, T. and Frederiksen, P., 2002. Change detection for automated updates of the TOP10DK map database using aerial photos. In: K. Åstrøm (ed.), Proceedings SSAB'02, Symposium on Image Analysis, Lund, March 7-8, 2002, Lund University, Centrum Scientarum Mathematicarum, pp. 61–64.

Smith, S. and Brady, J., 1997. SUSAN—a new approach to low level image processing. International Journal of Computer Vision 23(1), pp. 45–78.

Vincent, L. and Soille, P., 1991. Watersheds in digital spaces: An efficient algorithm based on immersion simulations. IEEE Transactions on Pattern Analysis and Machine Intelligence 13(6), pp. 583–598.