FEATURE EXTRACTION FOR QUALITY ASSESSMENT OF AERIAL IMAGE SEGMENTATION

V. Letournel[a,b,⋆], B. Sankur[c], F. Pradeilles[b] and H. Maître[a]

[a] Dept. TSI, ENST, 75634 Paris Cedex 13, France - letour@tsi.enst.fr

[b] DGA/Centre Technique d'Arcueil, 16 bis av. Prieur de la côte d'or, 94114 Arcueil Cedex, France - valerie.letournel@etca.fr

[c] Boğaziçi University, 80815 Bebek Istanbul, Turkey - sankur@boun.edu.tr

KEY WORDS: Protocol, Performance, Segmentation, Interpretation, Features, IKONOS, Statistics

ABSTRACT

We present a new evaluation methodology and a feature extraction scheme for segmentation algorithms in the context of photo-interpretation. The novelty of the proposed methodology is that subjective evaluation marks are involved in the determination of the feature subspace. In fact, our aim is to determine features in alignment with the perception of photo-interpreters, alternatively called psychovisual features. The proposed methodology was applied to the detection of building targets in aerial images. More specifically we considered the delineation of polygonal buildings in semi-urban areas on IKONOS images (1 meter resolution). We determined from the images, concurrently, various objective performance measures and collected votes of a jury of evaluators. The methodology to find the concordance between objective features and subjective marks was the canonical analysis of tables.

1 INTRODUCTION

A plethora of image segmentation algorithms have been advanced in the last decades, and their variety is still on the increase. There is therefore a urgent need for techniques to assess objectively the merits and performance advantages of these algorithms in the context of various vision tasks. A seminal work in this direction is the method developed by Zhang (Zhang, 1996), which is based on the accuracy of feature measurements of the segmented objects.

In the taxonomy of methods for the evaluation of segmentation algorithms several approaches can be distinguished. One class of a priori methods (Ji and Haralick, 1999) try to predict the algorithmic performance vis-à-vis generic inputs, before any implementation. Another class of a posteriori methods need the actual output of algorithm, and use, in the absence of ground-truth reference, 'goodness of segmentation' measures (Huang and Dom, 1995). These measures are based on the characterization of the outcome, such as, the consistency of features within the segments, smoothness along the contours or high contrast across the boundaries. However the most common evaluation method in the literature relies on the discrepancy measures as in (Kanungo and Haralick, 1995), that is, the differences between an ideal segmentation map, called the "ground-truth" and the actual segmentation outcome. The typical difference criteria are missed object pixels, false alarm pixels localization errors, mismatch of edges, shape discrepancy etc.

It is more relevant to evaluate the usefulness of an image segmentation algorithm in the context of a specific task rather than try to address the general segmentation performance issue. A case in point is the photo-interpretation of aerial images where we want to assess how much specific algorithms and/or features aid in the completion of vision tasks. In such vision tasks as target detection, battle damage assessment, delineation of buildings and man-made objects, the segmentation map represents an intermediate level intelligence to the human operators. It is then necessary that the delineation of the objects and the features emphasized be in concordance with the expectation of the photo-interpreters, hence satisfy human vision requirements.

In this work we limit our task to the extraction of buildings in aerial images. We first explore the relevant features that characterize buildings in aerial images, with the ultimate goal of identifying the "psycho-visual" features, which are largely correlated with the photo-interpreters' attention mechanisms. In other words we introduce a perceptual dimension when we evaluate the performance of segmentation algorithms in terms of "what the photo-interpreters prefer and judge as relevant".

The organization of the paper is as follows. In Section 2 we explain the framework of application and the motivation for a segmentation evaluation methodology, where humans are in the loop. The proposed methodology is detailed in Section 3. Results of the selected features and their validation against the feedback received from photo-interpreters are detailed in Section 5. Finally Section 6 draws the conclusions.

2 PROBLEM STATEMENT

Interpretation and annotation of aerial images is an important task in various military and non-military contexts. We intend to establish the qualifications of a segmentation algorithm judged to be an effective tool by the photo-interpreters. Since the algorithms are qualified according to their goodness-of-segmentation features we have received feedback from photo-interpreters in terms of their subjective quality judgements.

We aim to assess segmentation algorithms based on the understanding of the reasoning of photo-experts, that will mimic the human judgment on the quality of an extracted object. In other words the similarity of the two objects, the ground-truth object in the scene and the actual extracted object, will be based on human similarity assessment. For instance, human judgement is more sensitive to a false sharp protuberance from an object, albeit small in pixel count, than to that object's dilation, so that the simple total

of misclassified pixels may not be after all a good quality measure. We expect that the difference score attributed by the evaluation algorithm must well reflect the subjectively perceived differences. Obviously a simple pixel-to-pixel comparison of the experimental and ground-truth segmentation maps may prove very inadequate, since they are devoid of operator's perception, between the ground-truth and the actual object.

2.1 The Tverskian approach

A measure which takes subjective assessment of similarity into account is the "feature contrast model" first proposed by Tversky (Santini and Jain, 1999). In the Tverskian approach, objects are characterized by a set of binary attributes, and (dis)similarity is measured in the attribute space, relying on the notion of psycho-visual similarity. Tversky considers separately the effect of matching features between objects as well as the aspects in which they differ. They are represented by binary values so that stimuli of the perceiver are characterized by the presence or absence of these features.

However it is cumbersome to represent such numerical (non-categorical) values. Furthermore in computer vision one cannot usually obtain binary features due to noise in measurements. This has lead Santini *et al.* (Santini and Jain, 1999) to introduce the use of fuzzy predicates in the contrast model. So the similarity between two fuzzy sets Φ et Ψ corresponding to measurements made on two images (for example Φ : measurements on the ground-truth and Ψ on the segmentation result) is expressed as :

$$\begin{aligned} S(\Phi,\Psi) &= \sum_{i=1}^{p} \min\{\mu_i(\Phi),\mu_i(\Psi)\} \\ &\quad -\alpha \sum_{i=1}^{p} \max\{\mu_i(\Phi)-\mu_i(\Psi),0\} \\ &\quad -\beta \sum_{i=1}^{p} \max\{\mu_i(\Psi)-\mu_i(\Phi),0\} \quad (1) \end{aligned}$$

where μ_i is the membership function of the ith predicate, and p is the number of predicates measured on images.

2.2 The proposed approach

In an effort to establish segmentation features relevant to the human judgement we use the method of canonical analysis of tables between two feature spaces. One feature space consists of objective features of the segmented "building" object; the other feature space consists of subjective features on the same object, expressed categorical quality points given by people. We transform one space toward the other to render them as "parallel" as possible. The degree of parallelism achieved is a measure of the relevance of feature set combinations to the human judgement. Presently our approach is not based on fuzzy membership functions as in (Santini and Jain, 1999) but on predictability of one set of variables (subjective votes) by the another set of variables (objective features). Thus the similarity measure is a linear combination of feature values.

We deal with IKONOS images (1 meter resolution) of semi-urban areas and the task to delineate polygonal buildings. Two categories of objective features are considered : features specific to the geometry of buildings, called intrinsic features, and features related to the appearance models, that is the gray-level contextual information, called extrinsic features. In the first set, we focus on form and size (a rectangular and/or big blob is more significant than a small and/or non-rectangular one), parallelism of the opposite sides, number of corners, regularity of edges, (e.g., a closed and smooth contour is more significant). In the second set we pay attention to the shadow effects near the edges, gray-level uniformity inside and contrast with immediate surrounding etc..

3 FEATURES OF BUILDINGS IN AERIAL IMAGES

The set of intrinsic and extrinsic segmented building features we consider are listed below. In what follows Z will denote a segmented generic building region.

3.0.1 The intrinsic features These features are measured vis-à-vis the ground-truth data :

1. The elongation index, $IA(Z)$ (Coster and Chermant, 1985), of a segmented region (Z) is defined as :

$$IA(Z) = \frac{\pi L_g^2(Z)}{4A(Z)} \quad (2)$$

where $L_g(Z)$ is the geodesic diameter of Z and $A(Z)$ is its area. Note that for a disk its value is minimum and equal to 1. The elongation index can be instrumental in distinguishing, for example, a "U form" from a rectangle having the same perimeter and area. A case in point would be a multi-winged building.

2. The compactness $C(Z)$ defined by :

$$C(Z) = \frac{4\pi A(Z)}{p^2(Z)} \quad (3)$$

where $p(Z)$ is the perimeter of the boundary of Z. Recall that $C(Z) = 1$ for a disk, and goes to zero for very elongated forms or regions with severely jagged edges.

3. The bounding box $BB(Z)$ (Coster and Chermant, 1985) is constructed along the inertial directions of the extracted region. Two features have been extracted indicating the degree of rectangularity of the region. The first one is the excess difference and the second one is the excess ratio, respectively, of the building pixels and of its bounding-box pixels :

$$D(Z) = A(BB(Z)) - A(Z) \quad (4)$$

$$R(Z) = \frac{A(Z)}{A(BB(Z))} \quad (5)$$

4. Object symmetry: Man-made objects, like buildings, have usually strong symmetry property. We adapt a measure introduced in (Colliot et al., 2002) by computing the symmetry score for different positions of the symmetry axis (Π) passing through the centroid :

$$0 \leq S(Z, e_\Pi(Z)) = \frac{|Z \cap e_\Pi(Z)|}{|Z \cup e_\Pi(Z)|} \leq 1 \qquad (6)$$

where $|.|$ denotes the cardinality of a set, and $e_\Pi(Z)$ is the mirror reflection of the set Z with respect to the Π axis. The measure counts the number of pixels that have a symmetric counterpart with respect to the Π axis. One searches for the optimum orientation of the axis, which corresponds to the position where the maximum number of pixels have their matching counterparts across the Π axis. For a rectangle two maxima are found located at θ_1 and θ_2 and with amplitude A_1 and A_2. Then we derive two measures :

$$ms_1 = \frac{1}{2} \sum_{i=1}^{2} (1 - |A_{ref,i} - A_i|) * \exp(-\alpha * |\theta_{ref,i} - \theta_i|) \qquad (7)$$

$$ms_2 = \frac{1}{2} \sum_{i=1}^{2} (1 - |A_{ref,i} - A_i|) * (1 - \alpha * \sin^2(\theta_{ref,i} - \theta_i))^+ \qquad (8)$$

where ref denotes the ground-truth values, and $^+$ is the positive part of the function.

5. Histogram differences : We expect the histogram of a correctly segmented object to follow very closely that of its ground-truth object histogram. Low resolution histograms, H, with only 16 gray levels, were calculated since the data available from small objects may be very limited. The discrepancy between the gray-level histograms is estimated by using the χ^2 and L_2 metrics (Erdem et al., 2001), normalized to the range $[0, 1]$:

$$d_{\chi^2}(H_1, H_2) = \frac{\sum_{j=1}^{b} \frac{[R_1 H_1(j) - R_2 H_2(j)]^2}{H_1(j) + H_2(j)}}{N_{H_1} + N_{H_2}} \qquad (9)$$

$$d_{L2}(H_1, H_2) = \sqrt{\frac{\sum_{j=1}^{b} [R_1 H_1(j) - R_2 H_2(j)]^2}{NS_{H_1} + NS_{H_2}}} \qquad (10)$$

where $b = 16$ denotes the number of bins in the histogram, the scaling parameters R_1 and R_2 are used to normalize the data when the total number of elements in the two histograms are different, $N_{H_i} = \sum_{j=1}^{B} H_i(j)$, $NS_{H_i} = \sum_{j=1}^{B} H_i^2(j)$, $R_1 = \sqrt{\frac{N_{H_2}}{N_{H_1}}}$ and $R_2 = 1/R_2$.

3.0.2 The extrinsic features The extrinsic feature set deals with appearance of the object, and not with its ground-truthed geometric characteristics. They penalize or reward the generic goodness of the segmented region.

1. The uniformity and contrast of the segment: We intuitively expect that a segmented object be "more uniform" as compared to "its surroundings". We obtain the surrounding region $\bar{Z}$ as the dilation of the object's bounding box, shown in figure 1 (pixels in white). The notions of "object uniformity" and of "object contrast" are quantified as follows

$$d_{u1} = \frac{\sigma_{\bar{Z}}^2}{\sigma_Z^2} \qquad (11)$$

$$d_{u2} = \frac{|\mu_Z - \mu_{\bar{Z}}|}{\sigma_{\bar{Z}} \sigma_Z} \qquad (12)$$

where μ_Z and σ_Z^2 are the mean and variance of the object.

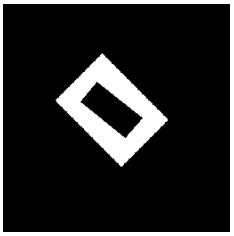

Figure 1: A segmented form (rectangle in black) and its surround (in white) obtained by dilating its bounding box.

2. Contour regularity: We define this notion as the mean absolute curvature as in (Chassery and Montanvert, 1991) :

$$C_4 = \frac{1}{p} \sum_{i=1}^{N} |c_{4,i}| \qquad (13)$$

where p is the perimeter of the boundary and $c_{4,i}$ is the fourth order curvature (see figure 2) of the ith edge pixel.

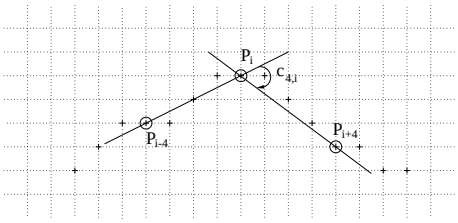

Figure 2: Curvature ($c_{4,i}$) of point P_i defined as the angle formed by the line (P_{i-4} P_i) and (P_i P_{i+4}) (P_{i-4}, P_i and P_{i+4} belong to the boundary).

3. Object contrast: Well segmented objects must have distinct gray levels with respect to the background. In the definition of contrast, given in (Erdem et al., 2001) one computes the mean grey level over blocks 'just inside' ($\bar{N}_Z^j$) and 'just outside' ($\bar{N}_{\bar{Z}}^j$) for j points regularly spaced along the boundary. These blocks, typically 3x3 or 5x5, 'just inside' and 'just outside' are drawn on the two sides of the j normal to the boundary.

$$d_c = 1 - \frac{1}{n} \sum_{j=1}^{n} \frac{|\bar{N}_Z^j - \bar{N}_{\bar{Z}}^j|}{255} \qquad (14)$$

where n is the number of normals.

4 METHODOLOGY

In order to select features of buildings in aerial images that are both statistically discriminating, and at the same time judged relevant by photo-interpreters, we first build a ground-truthed segmented image database. Then we obtain the subjective scores by having evaluators to vote on their quality. The evaluators view the displayed segmentation outcome from algorithms side by side with their originals. Concurrently we extract from the segmented images and their ground-truths objective performance scores based on feature differences as in Section 3. Finally we carry out a canonical analysis of subjective and objective quality scores in order to obtain the best possible match between these two tables and thus determine the relevant "psychovisual features". The details of the proposed methodology are described in the following paragraphs.

4.1 Construction of the Segmented Image Database

We have selected four characteristic segmentation algorithms. These belong to the segmentation paradigms based on image discontinuity, image similarity and feature-space clustering. Several varieties of each algorithm were obtained by adjusting their parameters. We used 1) a split-and-merge algorithm by Suk (Suk and Chung, 1983) (three parameters to be set) ; 2) the Canny-Deriche edge detector followed by hysteresis thresholding and edge closing (four parameters involved) ; 3) a feature-space algorithm that uses watersheds of the image histogram, smoothed by a multi-fractal measure (Kam, 2000) (four thresholds) ; 4) an image similarity algorithm, the seeded region-growing algorithm (Gagalowicz and Monga, 1985) (one parameter involved).

As image material, we have chosen nine sub-images from an IKONOS image (1 meter resolution) of the area of Algiers. The scenes are rich with polygonal buildings. For each image we have established the ground truth by manual tracing with a photo-interpreter tool. Using different settings of the parameters of the above segmentation algorithms we obtained in total 160 segmentation outcomes. Seven of nine images have been segmented, each, with 20 variations of the segmentation algorithms while two of them with only 10 versions. This gave us a total of 160 (= 20*7 + 10*2) segmented scenes to be voted on (Letournel, 2000). In the sequel we will refer to the segmentation result obtained with a given algorithm and a given parameter setting simply as "segmentation".

4.2 Subjective segmentation measures

A group of subjects evaluated the set of segmented images and gave their assessment marks. The marks were in the [-2, 2] range, going from the lowest quality mark of '-2' as "unacceptable" to '+2' meaning "near perfect". The subjects could view side by side the segmented test image and its "perfect" ground-truth segmented version. An instance of the test image is shown in figure 3 (edges are in white) and its reference image in figure 4. To avoid any fatigue effects on the voters we decided to partition the segmentation database of 160 images into 4 groups of 40 images. Each voter was randomly assigned to one of these 4 groups. We made sure that groups are formed of a fair distribution of "good" and "bad" segmentations.

Figure 3: A segmented image to be marked (edges are presented in white).

Figure 4: Ground-truth segmentation used as reference to mark segmentation on figure 3 (edges are in white).

4.3 Elaboration of the features space

We have first used the principal component analysis (PCA) on the subjective features, with the goal of ascertaining the coherence among the evaluators. Secondly we have applied PCA to the objective features to understand if there would be a more appropriate subspace describing them. Finally we studied the two sets (objective and subjective features) jointly using Canonical Analysis (CA). Recall that the aim of this particular statistical tool (CA) (Saporta, 1990) is to put into evidence any linear relationship that may exist between two sets of quantitative measurements on the same events.

More formally when n events are described by two sets of variables (respectively, of dimension p and q), one searches for a linear combination of variables of set 1 (P) and a linear combination of variables of set 2 (C) that are the most correlated with each other. In our context these two sets are obviously the set of objective features measured (P) and the set of marks given by evaluators (C). The n observations consist of the 40 segmentations in each group, each taking place in $\mathbb{R}^n$. Notice that the CA is run separately for each one of the four groups. Let's denote by $\mathbf{x^1}, ..., \mathbf{x^P}$ the objective feature measurements, where p = 12 and component $\mathbf{x^i}$ is the ith feature. All this data is organized in

a nxp matrix X. Similarly let $\mathbf{y^1}, ..., \mathbf{y^q}$ denote the subjective features, representing the marks of the q evaluators, where $q = 10$, when we also take the photo-interpreters votes into account. All the subjective data is organized in a nxq matrix Y.

To compare the two sets, we calculate linear combinations of the measurements in set 1 and set 2 :

$$\xi = X\mathbf{a} = a_1\mathbf{x^1} + ... + a_p\mathbf{x^p} \quad (15)$$

$$\eta = Y\mathbf{b} = b_1\mathbf{y^1} + ... + b_q\mathbf{y^q} \quad (16)$$

where the projection vectors ($\mathbf{a}$ and $\mathbf{b}$) are to be determined by maximizing the squared *canonical* correlation between ξ and η under the constraint that ξ and η are unit norm vectors. ξ and η are called *canonical features*. If the matrices $X'X$ and $Y'Y$ are invertible, one finds that $\mathbf{b}$ is the eigenvector of the matrix $M = (Y'Y)^{-1}Y'X(X'X)^{-1}X'Y$ related to the biggest eigenvalue, β_1^2. Similarly, $\mathbf{a}$ is the eigenvector of $N = (X'X)^{-1}X'Y(Y'Y)^{-1}Y'X$ associated with the same eigenvalue. When the first couple of canonical features (ξ_1,η_1) has been found, we proceed then to the following couple (ξ_2,η_2) so that their correlation is the next largest in order, while at the same time ξ_1 and ξ_2 (respectively η_1 and η_2) have zero cross-correlation, and so on for ξ_3, η_3 ... From a geometrical point of view, on finds that (β_1) is the cosine of the smallest angle between spaces, respectively generated by the columns of X and Y. The canonical analysis problem can be compared to the problem of multiple regression (for more details see (Saporta, 1990)).

To choose r features among the p calculated, we used the redundancy criterion proposed by Thorndike (Tinsley and Brown, 2000). Let's call as *intraset loadings* the correlations of the observed variables in set P with its canonical features and as *interset loadings* the correlations of the observed variables in set P on the canonical components of set C. Drawing on principal component thinking, we can set an analogy between the eigenvalues of PCA and the sum of squared intraset loadings of the variables on a canonical component. The latter is the amount of variance of the set that is accounted for by the canonical variate of that set. This quantity, divided by the number of variables, produces the proportion of variance in the set that is accounted for by the component, denoted V_{C_j}, for the j^{th} component of set C. Let's recall at this moment that the squared canonical correlation (β_j^2) is the proportion of a canonical component's variance accounted for by the paired component in the other set. Therefore the proportion of variance in set P accounted for by the j^{th} component of set C is : $Red_{P_j} = V_{P_j} * \beta_j^2$.

5 RESULTS

5.1 PCA of the subjective features

We formed a heterogeneous jury of 32 evaluators differing in their expertise and familiarity with images. Some of them were not from image processing field (like secretaries ...), others are doctoral students, technicians, professors, yet others were infrared or radar images domain experts. In addition each of 4 photo-interpreters have marked 2 groups of images. In the final analysis each image has ended up receiving 10 marks. They were summed up in 4 matrices (corresponding to the four groups defined in Section 4.1) of 40x8 elements. An element y_i^j corresponds to the mark given by the jth evaluator ($j = 1, ..., 8$) to the ith image ($i = 1, ..., 40$).

Since not every image was marked by every evaluator, the PCA on the subjective features had to be carried out in groups. We have observed that the first PCA axis carries about 80 % of the inertia for all groups. It represents, in fact, the baseline of consensus of the evaluators. We should note that the data from the photo-interpreters has not taken place in this computation, but their mark vectors have been subsequently projected on the principal axes for validation and are also highly correlated with the first axis. The second component can be interpreted as portraying the differential behavior of the evaluators, that is their tendency for severe or tolerant voting (a "severe" evaluator gives much easily a negative mark). In figure 5 we show the coordinate axes $u1, u2$ of the first two largest eigenvectors of group 1, and the projection of the group's mark vectors. All three groups 1, 3, 4 have similar projections, whereas the group 2 behavior differs on the vertical axis. One explanation could be that this group 2 has a relatively larger proportion of inexperienced evaluators. Indeed the second component axis for this group seems only differentiating inexperienced and experienced evaluators. Let's note that the principal axes do not have necessarily the same interpretation from group to group.

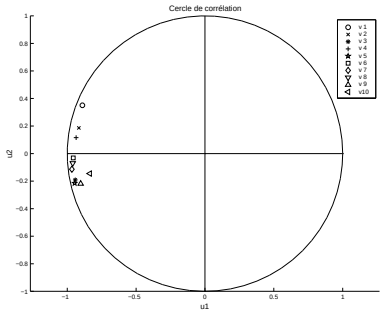

Figure 5: Projection of mark vectors of the 8 evaluators (v1 to v8) and the 2 experts (v9 and v10) for group 1 on the principal PCA axes.

5.2 Canonical analysis of the objective and subjective features

As in the preceding PCA case of subjective scores, we implement the canonical analysis per group. Our goal is to find the subset of features which is the most correlated with the baseline of consensus of the evaluators found with the PCA.

As we do not know a priori the cardinality r of the subset, we carry out an exhaustive canonical analysis. For each value $r \in [1, p]$, we try all possible combinations of r features chosen among p, and we run the CA on this subset

and C, and we determine the canonical feature η_j which maximizes V_{C_j} (and thus represents the baseline of consensus) and the redundancy Red_{P_j}. We keep the subset of features which maximizes Red_{P_j}, among all the possible combinations (the process takes less than 15 minutes on a 333 MHz processor). We show for each group on table 1

	r	Features	Red_{P_j}
G1	5	Ia, C , C_4, m_{s1}, d_{L2}	0.59
G2	7	Ia, D, R, d_{u1}, C_4, m_{s1}, d_{χ^2}	0.37
G3	10	Ia, D, R, d_{u1}, d_{u2}, C_4, d_c, m_{s1}, d_{χ^2}, d_{L2}	0.68
G4	8	Ia, C, D, R, C_4, d_c, d_{χ^2}, d_{L2}	0.64

Table 1: Feature subsets which maximize redundancy.

the feature subset that maximizes the redundancy and the value of the redundancy criterion. Groups 3 and 4 have the highest scores, so one can claim that the subjective and objective features are substantially related along the first canonical dimension. But this extent of agreement between numerical features and evaluators cannot be reached for group 2, which shows quite a low redundancy score.

When we compare the behavior of the mean of the marks (recall that the PCA has revealed a 1-dimensional evaluator space) and that of the canonical features, we observe that some segmentations cause great discrepancy. In simpler words for these images the note of the evaluators cannot be predicted based on the chosen feature set. Thus we developed a procedure to eliminate segmentations that cause conflicting votes among the evaluators and kept only the consensus images, that is those images that received the same relative ranking from all the evaluators. CA results with the pruned consensus set are shown in table 2. This second table reveals a great improvement both in the

	r	Features	Red_{P_j}
G1	7	Ia, C , D, R, d_{u2}, C_4,et d_c	0.73
G2	6	Ia, d_{u1}, C_4,et m_{s1}, d_{χ^2} et d_{l2}	0.40
G3	10	Ia, C , D, R, d_{u1}, d_{u2}, C_4, m_{s2}, d_{χ^2} et d_{L2}	0.79
G4	7	Ia, d_{u1}, d_{u2}, C_4, d_c, d_{χ^2} et d_{L2}	0.80

Table 2: Feature subsets which maximize redundancy (learning on images of consensus).

consistency between the selected features across groups as well as in the redundancy marks. While groups 1, 3 and 4 show this improvement, group 2 remains still a poor predictor of evaluator marks from the features. Hence we removed this group from the rest of the experiment.

Thus we get three psychovisual feature subsets which predict the vote of the evaluators reasonably well. One method to collapse these three sets to one "best" set would be to use cross-validation across groups. For example we use the feature set of group 1 and use it on predicting the data of the other two groups, i.e., groups 3 and 4, and calculate the redundancies $Red(Ia, C, D, R, d_{u2}, C_4, d_c)$ on these groups. We repeat this calculation similarly for the other two feature sets. Then we take the average of the redundancies of each feature set on the three segmentation groups, and choose the largest one.

6 CONCLUSION

We have presented the framework for a new feature extraction method for a task-oriented segmentation that combines both the statistical properties of image features and segmentation quality assessments of a jury. The methodology has been applied to the task of segmentation buildings in medium-resolution images. This study was the first step for the extraction of psychovisual image features. The work will continue to build membership functions of features in a Tverskian context.

ACKNOWLEDGMENTS

We would like to thank Pr. B. Burtschy (from ENST) for his valuable advice during discussions on canonical analysis.

REFERENCES

Chassery, J. M. and Montanvert, A. (1991). *Géométrie discrète en analyse d'images*. Hermès, 34, rue Eugène Flachat 75017 Paris.

Colliot, O., Bloch, I., and Tuzikof, A. V. (2002). Characterization of approximate plan symmetries for 3d fuzzy objects. In *IPMU, Annecy, France*.

Coster, M. and Chermant, J. (1985). *Précis danalyse d'images*. CNRS.

Erdem, C., Tekalp, A., and Sankur, B. (2001). Metrics for performance evaluation of video object segmentation and tracking without groundtruth. In IEEE, editor, *International Conference on Image Processing (ICIP'01)*, volume 2, pages 69–72.

Gagalowicz, A. and Monga, O. (1985). Un algorithme de segmentation hiérarchique. In INRIA, editor, *Reconnaissance des Formes et Intelligence Artificielle (Grenoble)*, volume 1, pages 163–178.

Huang, Q. and Dom, B. (1995). Quantitative methods of evaluating image segmentation. In IEEE, editor, *International Conference on Image Processing (ICIP'95), Washington DC, USA*.

Ji, Q. and Haralick, R. M. (1999). Quantitative evaluation of edge detectors using the minimum kernel variance criterion. In IEEE, editor, *International Conference on Image Processing (ICIP'99), Kobe, Japan*.

Kam, L. (2000). *Approximation multifractale guidée par la reconnaissance*. PhD thesis, Orsay University.

Kanungo, T. and Haralick, R. M. (1995). Receiver operating curves and optimal bayesian operating points. In IEEE, editor, *International Conference on Image Processing (ICIP'95), Washington DC, USA*, volume 3, pages 256–259.

Letournel, V. (2000). Avancement de thèse. Technical Report CTA 2000 R 086, Centre Technique d'Arcueil.

Santini, S. and Jain, R. (1999). Similarity measures. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(9):871–883.

Saporta, G. (1990). *Probabilités, analyse des données et statistique.* Technip.

Suk, M. and Chung, S. M. (1983). A new image segmentation technique based on partition mode test. *Pattern Recognition*, 16(5):469–480.

Tinsley, H. E. A. and Brown, S. D. (2000). *Handbook of Applied Multivariate Statistics and Mathematical Modeling.* Academic Press.

Zhang, Y. J. (1996). A survey on evaluation methods for image segmentation. *Pattern Recognition*, 29(8):1335–1346.