LICENSE-PLATE LOCALIZATION BY USING VECTOR QUANTIZATION

Stefano Rovetta and Rodolfo Zunino

DIBE - Dept. of Biophysical and Electronic Engineering - University of Genoa
Via all'Opera Pia 11a - 16145 Genova - Italy

ABSTRACT

The paper describes a novel approach using Vector Quantization (VQ) to process vehicle images for automated identification. The VQ-based method yields superior quality in picture compression for archival purposes, and, at the same time, supports the localization of text regions in the image effectively. As opposed to standard approaches, VQ encoding gives some hint about the contents of image regions; such information is exploited to boost localization performance. The VQ system may be trained empirically from examples; this provides adaptiveness and on-field tuning facility. The approach has been tested in a real application and included satisfactorily into a complete system for vehicle identification.

1. INTRODUCTION

In the area of vehicle identification for Automated Transport Systems (ATS), visual recognition of license plates is currently the most promising approach, as passive sensing prevents vehicles from carrying on-board transponders. On the other hand, the high-dimensional signals involved by such methods impose a notable computational load to an automated system.

In a visual vehicle-identification system, low-level imaging first restores signal quality by application-specific techniques (e.g., filtering, contrast enhancement, etc). In a second step, the localization of interesting scene regions is attained by a segmentation process, involving classical image-processing algorithms [1], genetic-algorithms techniques [2], and neural-network models [3,4]. Localization prepares actual vehicle identification, usually supported by Optical Character Recognition (OCR) methods. In applications requiring archival facility for time-logging purposes, image compression algorithms also apply to minimize storage space.

This paper tackles the problem of license-plate localization in visual signals, and presents a novel methodolgy that exploits Vector Quantization (VQ) [5] as the basic image-processing paradigm. As compared with other approaches, VQ enables a system to support both plate localization and image archival simultaneously and efficiently. The baseline of the present work is that using VQ in image coding yields compression ratios and visual quality that outperform standard methods. The described research shows that a specific use of VQ can make localization straightforward. A theoretical framework proves that VQ image representation matches the operational requirements of the localization process. The VQ-based localization/archival method has been tested in a complete automated system for car-ID applications. The approach proved very satisfactory as to both localization accuracy and coding effectiveness.

Section II first describes briefly the core of VQ-based image representation and coding; then the theoretical justification of VQ-based localization is given. Section III presents practical algorithms to set up a localization system, whereas Section IV reports on experimental results. Some concluding remarks are made in Section V.

2. GENERAL FRAMEWORK FOR VQ

2.1 VQ-based image coding

Compression is the reference application area of Vector Quantization [5] in 2-D signal processing [6]. In the specific application, the localization process requires VQ to operate in the pixel domain (rather than in the usual frequency domain). In VQ-based image coding, input images are split into elementary blocks. Such blocks span vectors in a data space, where the quantization process defines a predetermined, fixed "codebook" of reference vectors ("codewords"). The coding process associates each block with the codeword that optimizes a similarity criterion, and encodes the block by the codeword's index. Euclidean distance usually measures distortion in block-codeword matching. Compression derives from using a codebook that is "small" as compared with the number of possible blocks. Therefore, image coding is a lossy process, as reconstructed images differ from the original ones due to quantization noise.

The algorithm placing codewords in the data space ultimately determines the compression ratio and the overall distortion of a coding system [7,8]. VQ training methods use an example-driven strategy: the algorithm is supplied with a set of image blocks, and outputs the set of codewords minimizing distortion over training samples. The possibility of sample-driven training represents a crucial advantage of VQ, which makes a system application-adaptive and greatly improves flexibility. The research presented here adopted a novel algorithm specifically designed to minimize codebook size without affecting reconstruction quality [9].

Some technical improvements augment the basic schema. *Mean Residual Coding* (MRC) subtracts a block's mean value from the block's pixels before VQ encoding. This makes the block-coding process brightness-independent. *Variable block size*, on the other hand, enhances compression: uniform image regions may be covered by larger blocks, whereas detail-rich portions require smaller blocks to render visual information properly. Each block's size is typically set by a threshold mechanism measuring the variance of a block's pixels. If square blocks are used, each

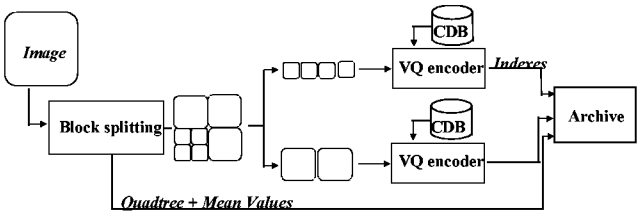

Fig.1 – VQ-based image-compression schema

block-splitting generates four equal subblocks, hence a quadtree minimizes the consequent structural representation.

The overall result of a VQ-based coding session (Fig.1) includes three data structures: a quadtree ruling block layout, a set of mean values giving the blocks' average brightness, and a set of codeword indexes to refine details within each block.

2.2 Using VQ for license-plate localization

The quantization principle can greatly facilitate the localization process because the coding process involves an implicit analysis of the image contents. A codebook is defined in the same (pixel) space of the encoded blocks, hence associating each block with the best-matching codeword implies a classification of the block contents. Classification results may give some hint about the block content itself, and in particular, whether the block is likely to cover a license plate. The overall localization process can be formalized as follows: define the image-content function as

$$t(\mathbf{p})=\begin{cases}1 & \text{if pixel } \mathbf{p}\equiv(x,y) \text{ belongs to a license plate}\\ 0 & \text{otherwise}\end{cases}$$

Then the general problem of localization is equivalent to locating the image region, S*, that satisfies

$$S^{*}=\max_{x_0,y_0,x_1,y_1}\int_{x_0}^{x_1}\int_{y_0}^{y_1}t(x,y)\,dx\,dy \qquad (1)$$

where a license plate has been (realistically) assumed to be framed by a rectangular box. In the following, the rectangular regions considered for localization will be called "stripes".

In the block-coding method, a stripe's boundaries will lay along the grid of blocks; this means that, in the VQ-based approach, the integral (1) becomes a sum of contributions from each block rather than a summation over single pixels. Moreover, when considering the results of VQ block classification, the limited number of codewords raises an ambiguity problem: the same codeword may encode both "interesting" blocks and irrelevant ones. Thus a block's contribution becomes a probability value, and the VQ-based localization problem (1) can be rewritten as

$$S^{*}=\max_{S}\sum_{\mathbf{b}\in S}p(t=1|\mathbf{b}) \qquad (2)$$

where $\mathbf{b}$ indexes the codewords associated with the blocks in region S, whereas $p(t=1|\mathbf{b})$ denotes the average probability that a block's pixels contain plate information, given the fact that the block is coded by codeword $\mathbf{b}$. The problem of selecting regions, S, in (2) will be addressed in Section III.

In order to attain VQ-based localization, each codeword is associated a "score", which estimates the corresponding probability of conveying plate information. Summing up (as per (2)) codeword contributions for each stripe gives the probability that the stripe contains the license plate. The highest-score image region is the localization output.

2.3 VQ-based training of codeword scores

The conditional probabilities in (2) cannot be estimated directly from a training set: the actual distribution of $\mathbf{b}$, covering all occurences of a codeword $\mathbf{b}$ in all possible images, is unknown. However, one can use Bayes' theorem and rewrite the conditional probability as

$$p(t|\mathbf{b})=\frac{p(\mathbf{b}|t)\cdot p(t)}{p(\mathbf{b})} \qquad (3)$$

Substituting (3) into (2) and disregarding constant terms yields a new problem formulation:

$$S^{*}=\max_{S}\sum_{\mathbf{b}\in S}p(t=1|\mathbf{b})=\max_{S}\sum_{\mathbf{b}\in S}\frac{p(\mathbf{b}|t=1)}{p(\mathbf{b})} \qquad (4)$$

This reformulation makes empirical training viable. The denominator of each term in sum (4) is the overall probability of codeword $\mathbf{b}$, and can be evaluated from relative frequencies by counting the occurrences of $\mathbf{b}$. The numerator is the probability of using $\mathbf{b}$ when the encoded block is known to cover a license plate. To estimate this quantity, one extracts from each training image the portion holding the license plate, and marks the used codewords accordingly. Relative frequencies will again give the required estimates. The final practical implementation may also differentiate scoring terms, so that the codewords coding license-plate blocks (or irrelevant portions) get higher rewards (or penalties). Codeword scoring proceeds off-line *after* training the codebook for VQ compression. Each codeword is eventually augmented by a content-dependent parameter, giving the likelihood that the codeword contributes to code a license plate. The following Section will integrate such VQ-based scoring mechanism with the topological stripe-extraction process to accomplish localization.

3. VQ-BASED LOCALIZATION

3.1 Stripe extraction

Stripe selection extracts an arbitrary number of candidate stripes, S, to be considered in (2). Such process exploits block-size adaptiveness, assigning smaller blocks to detail-rich image regions. License plates belong to such class of regions, due to the high contrast between text and background; average background information can also drive stripe identification, as license plates have either a bright or a dark background. Thus the set of interesting image regions can be easily compiled by searching the image quadtree for contiguous areas mapped by small-size blocks. This criterion proves quite robust because the quadtree

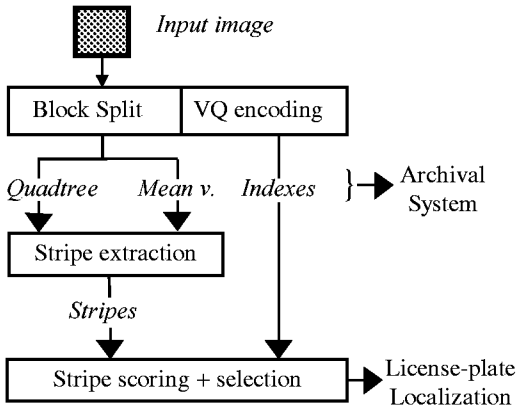

Fig.2 - The license-plate localization system

contrast-coding information is brightness independent. The only involved parameter is the expected stripe size. Stripe width and height do not exhibit high variance, as the eventual stripe size mostly depends on the relative sensor positioning (e.g., distance from the scene), which can be assumed to be controlled. Thus training the extraction module just requires to average size from sample stripes of license plates. Each extracted stripe is therefore scored by rewarding its consistency with the expected size; too long or too short stripes get a penalty term. Scored stripes then enter the actual localization process, which implements the strategy described in Section II.2.

3.2 Stripe scoring and license plate localization

The localization process first compiles the set of codewords used to encode the blocks in the extracted stripes. This makes it possible to retrieve the associated probabilities values; summing up such scores then makes it possible to label each stripe accordingly. The final score associated with a stripe results from two additive terms: the term from stripe extraction accounts for exterior stripe features (i.e., length), whereas the term from codeword scores takes into account expected region contents. Localization completes by selecting the highest-score stripe in the final sorted list. Figure 2 represents schematically the complete localization process, and points out the deep-rooted link with the VQ-based compression paradigm.

An important property of the schema is that the process need not issue a localization result for each input image: the system can also prompt a null output, such as "no valid stripes have been localized in the scene". This useful feature can be easily attained by setting a threshold on the stripe-selection mechanism: if no candidate stripe exhibits a satisfactory score, then no image region meets a reliability requirement for license-plate localization. Rejection ability is an advantage of using VQ to score a stripe's contents as opposed to "blind" approaches.

4. EXPERIMENTAL RESULTS

The localization mehodology has been tested in an Automated Transport System supporting vehicle identification, access grant, and billing in a parking area in Madrid. The system processes standard PAL input fields (interlaced half frames); the analog signal is converted into a digital picture holding 768×256 pixels with 256 gray-scale levels (8 bpp). Although the system does not de-interlace a complete frame for real-time constraints, both picture quality and localization performance are anyway satisfactory. A dedicated OCR method interprets license plates [10]. The current PC-based (C++) implementation runs under Windows NT. An image-understanding cycle (acquisition, localization, OCR) completes in about 200 ms on a standard Intel Pentium board running at 200Mhz. This performance proved satisfactory and cost-effective for the final application.

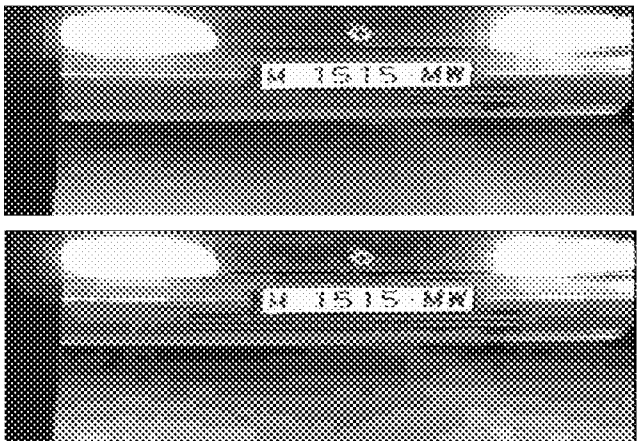

Fig.3 – Compression perform. *Cr* = 50; Top:VQ – Bottom: JPEG

The system's training process involved ten images; the codebook included 256 codewords for each block size and was trained with the algorithm presented in [9]. The codebook-fitting process completes in a few minutes on the PC architecture, hence optional on-line training can also be envisioned.

4.1 Image-compression performance

The relatively large size of pictures is the main reason for using VQ-based compression. The constraints on the image-coding method can be summarized as follows: 1) compression ratio must be large (Cr ≥ 40) to optimize archival storage; 2) the quality of reconstructed pictures must be acceptable. Experimental evidence shows that JPEG performance is unacceptable at high compressions; conversely, VQ outperforms other approaches at ratios Cr > 30. The comparison between the various approaches is beyond the scope of this paper; preliminary research [9] confirmed the advantage of VQ both quantitatively and qualitatively. The performances of the two coding schemata at high compression ratios (Cr=50) can be visually compared in Fig.3.

4.2 Localization performance

This section presents a complete example of the system's operation with intermediate results. Figure 4 (top) presents an input field and its quadtree decomposition (middle). The stripe-search process extracts candidate regions (Fig.4 – bottom); five candidates are identified and assigned a "spatial" score, rewarding how much they fit the ideal stripe length.

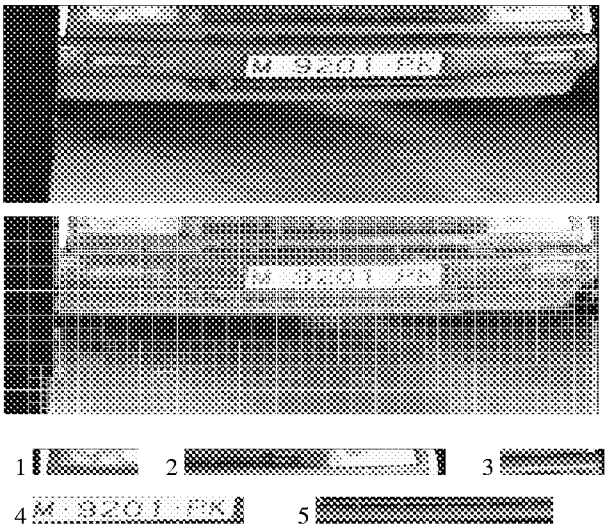

Fig.4 – Sample of the system operation
Top: input field – Middle: Structural decompositiion
Bottom: extracted candidate stripes

Table I. Results from the stripe-scoring systems in the example of Fig.3

	Spatial	*VQ*	*Final*
Stripe 1	-100	-1.66	-101.66
Stripe 2	0	-1.46	-1.46
Stripe 3	-100	-1.80	-101.80
Stripe 4	0	+0.99	+0.99
Stripe 5	0	-2.21	-2.21

The VQ-scoring system retrieves the codewords within each stripe and sums up the associated scores. The totals add to the spatial scores assigned by the stripe extractor, and end up with the final sorted list of candidates (Table I). The reliability of the "best" candidate is finally evaluated by the threshold-based rejection mechanism. In the present application the threshold was set to 0. In the example, stripe 4 is correctly prompted as the license-plate stripe.

The localization methodology underwent a thorough validation process in order to assess its performance on a statistical basis. The test set included more than 300 pictures, taken in different environmental conditions (e.g., brightness, sensor positioning, etc.). The OCR method adopted can handle two candidate regions simultaneously. Therefore, a localization process could be considered successful if the license plate appeared in the first two stripes of the sorted list; otherwise, a localization error was detected. Results indicate that the system performance exhibits an overall 2% error rate, which meets industrial requirements. Moreover, in the presence of an error, the OCR method adopted can detect the absence of text in a stripe and trigger exceptional processing accordingly. In practice, such critical pictures are filed in a special archive for delayed visual inspection.

5. SUMMARY

The crucial advantage of using a VQ-based representation is that the coding mechanism can give a localization system some hints about the contents of the candidate image regions. The core of the overall research, therefore, consists in best exploiting such information to boost localization performance. The possibility of example-driven on-field adjustment is a crucial advantage of the overall approach, as it gives a system flexibility and adaptiveness. Experimental results are satisfactory and the localization system integrated well with the surrounding ATS.

Future extensions address timing performance, and envision an electronic implementation with dedicated VLSI circuitry [11], or the adoption of acceleration techniques for codeword matching. These methods do not longer guarantee optimality in the best-matching search; hence the consequences of the speed/quality tradeoff are the object of current investigations.

6. REFERENCES

[1] Comelli P, Ferragina P, Granieri MN, Stabile F "Optical recognition of motor vehicle license plates" *IEEE Trans. on Vehicular Technology*, 1995, vol.44, No.4, pp.790-799

[2] Sang Kyoon K, Dae Wook K, Hang Joon K "A recognition of vehicle license plate using a genetic algorithm based segmentation" *Proc. IEEE Int.Conf. Image Proc.*, 1996, vol.2, pp.661-664

[3] Draghici S "A neural network based artificial vision system for license plate recognition" *Int.J. of Neural Systems*, 1997, vol.8, No.1, pp.113-126

[4] Nijhuis JAG, Ter Brugge MH, et al. "Car license plate recognition with neural networks and fuzzy logic" *Proc. IEEE Int.Conf. on Neur.Netw.*, 1995, vol.V, pp.2232-2236

[5] Gray RM "Vector Quantization" *IEEE ASSP Mag.*, 1984, vol.1, No.2, pp.4-29

[6] Nasrabadi N, King R "Image Coding Using Vector Quantization: a Review" *IEEE Trans. on Communications*, pp. 957-971, Aug. 1988

[7] Linde Y, Buzo A, Gray RM "An Algorithm for vector quantizater design" *IEEE Trans. on Communications*, Vol. COM-28, pp. 84-95, Jan. 1980

[8] Kohonen T, *Self-organization and Associative Memory, 3rd Ed.*, 1989, Springer Verlag

[9] Ridella S, Rovetta S, Zunino R "Plastic algorithm for adaptive vector quantization" *Neural Computing and Applications*, 1998, vol.7, No.1, in press

[10] Ancona F, Colla AM, Rovetta S, Zunino R "Implementing probabilistic neural networks" *Neural Computing and Applications*, 1997, vol.5, pp. 152-159

[11] Ancona F, Zunino R "Concurrent VLSI architectures for vector quantization" *IEEE Symp. Circuits and Sys. ISCAS'97* - Hong Kong - 1997, vol.III, pp.2076-2079