# Non-Supervised Robust Visual Recognition of Colour images using Half-Quadratic Theory

Rozenn Dahyot[1], Pierre Charbonnier[1] and Fabrice Heitz[2]

[1] Laboratoire régional des Ponts et Chaussées de Strasbourg
[2] LSIIT UPRESA CNRS 7005 Université Strasbourg I

**Abstract.** In this paper, a robust pattern recognition system, using a view-based representation of colour images is described. Standard appearance based approaches are not robust to outliers, occlusions or segmentation errors. The approach proposed here relies on robust M-estimators, involving non-quadratic and possibly non-convex energy functions. To deal with the minimisation of non-convex functions in a deterministic framework, we introduce an estimation scheme relying on M-estimators used in continuation, from convex functions to hard redescending non-convex estimators. At each step of the robust estimation scheme, the non-quadratic criterion is minimized using the half-quadratic theory. This leads to a low cost weighted least squares algorithm, which is easy to implement. The proposed robust estimation scheme does not require any user interaction because the algorithm automatically estimates all necessary parameters. The method is illustrated on a road sign recognition problem. Experiments show significant improvements with respect to standard estimation schemes.

## 1 Introduction

View-based representation of objects [8][7] has recently received considerable attention. A popular approach is the eigenspace representation, which allows a substantial dimensionality reduction of the recognition problem. Eigenspaces methods involve a reconstruction procedure, which consists in projecting the unknown image on the eigenspace. Traditional least-squares estimation is sensitive to gross errors (outliers) that occur, for instance, when the object is partially occluded. In order to cope with outliers, it is possible to reformulate the reconstruction step as a robust estimation problem. Among the methods proposed in the context of robust statistics [5][13][12][6][11], M-estimators offer a good compromise between algorithmic complexity and outlier rejection capability. The use of M-estimators for eigenspace recognition was first proposed by Black [2].

The first contribution of this paper is to formulate M-estimation in the framework of the half-quadratic theory (section 2). This theory introduces an auxiliary variable which, in our case, can be interpreted as an outlier mask. This provides a natural linearization of the normal equations, and results in a - fast and easy to implement - weighted least squares algorithm.

Parameter estimation is an important point for image reconstruction but is seldom addressed in the literature. Most of the time, the tuning of parameters is let to the user. As a second contribution of this paper, we exploit results from robust statistics to propose, in section 3, an algorithm that automatically estimates the scale parameter, making the whole method data driven. We also extend the method to colour image recognition, taking colour components into account in a pixel-by-pixel fashion. This is explained in section 4.

Finally, we apply our method to hard redescending non-convex M-estimators, which provide better outlier rejection. Following the same idea as in the Graduated Non Convexity (GNC) [2] algorithm, we gradually introduce non-convexity by using three M-estimators in continuation. However, in contrast with the GNC, our approach, presented in section 5, does not require any prior knowledge on the amplitude of residuals.

This work results in a new, simple to implement and non-supervised robust recognition scheme for colour images, which is applied to road sign recognition (section 6).

## 2 Robust estimation using the half-quadratic theory

To simplify, we first consider the case of grey level images. The extension to RGB images is presented in section 4. The training images are arranged as $n$-dimensional vectors by lexicographic ordering, and are normalised [7]. A Principal Component Analysis (PCA) is then performed and only $t$ eigenvectors are retained to span the eigenspace $F$. Given a (normalised) unknown image $e$, eigenspace techniques perform its reconstruction on $F$, i.e. compute the best representative $e^*$ of $e$ as a linear combination of eigenvectors $e^* = \sum_{j=1}^{t} c_j U_j$, where $c_j$ is the $j^{th}$ unknown co-ordinate of $e^*$ on the $j^{th}$ eigenvector $U_j$. The residual on the $i^{th}$ pixel is defined by:

$$\epsilon_i = e_i - e_i^* \tag{1}$$

A standard method for the estimation of $e^*$ consists in minimising the quadratic norm $J_0 = \|\epsilon\|$. Geometrically, the solution $\hat{e^*}$ of the least squares estimation of $e$, is the orthogonal projection of $e$ onto the $t$-dimension subspace $F$. As it is well known, least squares estimation is sensitive to gross errors (outliers) produced, for instance, by occlusions [2]. M-estimators, involving non-quadratic and possibly non-convex energy functions, are naturally robust to outliers or gross errors. M-estimation leads to the minimisation of a robust norm $J_1$:

$$J_1(c) = \sum_{i=1}^{n} \rho(\epsilon_i) \tag{2}$$

To minimise $J_1$ in a deterministic framework, we propose to use the *Half-Quadratic Regularization* theory [4]. Under certain conditions on $\rho$ [3], the non-quadratic energy $J_1$ is transformed into an augmented energy by introducing an auxiliary variable $b$.

$$\min_c \left\{ J_1(c) = \sum_{i=1}^{n} \rho(\epsilon_i) \right\} = \min_c \min_b \left\{ J_1^\sharp(c,b) = \sum_{i=1}^{n} (b_i.\epsilon_i^2 + \beta(b_i)) \right\} \tag{3}$$

where $\beta$ is a function of $b_i$. $J_1^\sharp$ is *half-quadratic*, i.e.:

- When $b$ is fixed, $J_1^\sharp$ reduces to a weighted least-squares criterion, whose solution satisfies:

$$(U^T \cdot B \cdot U)c = U^T \cdot B \cdot e \tag{4}$$

  where $B = diag_{i=1\dots n}\{b_i\}$ . There are many possible numerical algorithms to solve (4) (the conjugate gradient algorithm [10] is used here).
- When $c$ is fixed, $J_1^\sharp$ becomes convex with respect to $b$. Moreover, it can be shown [3] that the explicit minimizer is given by $b_i = b(\epsilon_i) = \frac{\rho'(\epsilon_i)}{2 \cdot \epsilon_i} \quad \forall i$. Due to the properties of $\rho$ [3], $b_i$ is close to one when $\epsilon_i$ is small (inliers), and vanishes for large values of $\epsilon_i$ (outliers). Therefore, $b$ can be seen as an *outlier mask*, conceptually similar to the one defined in [2], excepted that:
  - The mask is not Boolean in our case : $b_i$ is a real between 0 and 1,
  - $b$ appears naturally in our formulation and participates to the minimisation process.

Given an initial guess $c_0$, we use the following alternate minimisation algorithm:

$$\begin{cases} \forall i \in \{1 \dots n\} \ \epsilon_i^{(m)} = e_i - \sum_{j=1}^{t} c_j^{(m)} \cdot U_{ij} \\ \forall i \in \{1 \dots n\} \ b_i^{(m+1)} = \frac{\rho'(\epsilon_i^{(m)})}{2 \cdot \epsilon_i^{(m)}} \\ (U^T \cdot B^{(m+1)} \cdot U) \cdot c^{(m+1)} = U^T \cdot B^{(m+1)} \cdot e \end{cases} \tag{5}$$

We can notice that algorithm (5) is similar to the Location Step With Modified Weights, proposed by Huber in the context of robust statistics ([5] p.183).

## 3 Scale estimation

The robust estimator defined by expression (2) usually depends on a *scale parameter* $\sigma$, that controls the point where the influence of outliers begins to decrease:

$$J_1(c,\sigma) = \sum_{i=1}^{n} \rho(\frac{\epsilon_i}{\sigma}) \tag{6}$$

Several methods have been proposed to adjust parameter $\sigma$ [13][5]: joint estimation of $(c,\sigma)$ or estimation of $\sigma$ once and keeping it fixed. When $\sigma$ is fixed, the minimisation can be performed using (5) and modifying the weights according to $b_i = b(\frac{\epsilon_i}{\sigma})$. In this following, we focus on the joint estimation of parameter $\sigma$ and of vector $c$. Let us consider replacing $J_1(c,\sigma)$ by $J_2(c,\sigma) = \frac{1}{n}\sum_{i=1}^{n}[\ln(\sigma) + \rho(\frac{\epsilon_i}{\sigma})]$. This expression is attractive because it is directly related to a probability density of the residuals of the form $\frac{1}{\sigma} \cdot g\left(\frac{\epsilon_i}{\sigma}\right)$ with

$g(x) = \exp(\rho(x))$ ([5]p.176). Unfortunately, $J_2$ is non-convex with respect to $\sigma$. Besides, Huber notices this estimator is generally not robust and proposes to minimise the following modified objective function ([5] p.175-176):

$$J_3(c, \sigma) = \frac{1}{n} \sum_{i=1}^{n} \left[ \rho\left(\frac{\epsilon_i}{\sigma}\right) + a \right] \cdot \sigma \tag{7}$$

where $\rho$ must be a convex functions, and $a$ corresponds to a tuning parameter. It is obvious from (7) that minimising $J_3$ with respect to $c$ for a fixed scale is equivalent to minimising $J_1$. Besides, since $\rho$ is convex and under the assumption that the residuals $\epsilon_i$ are linear with respect to the $c_j$'s (which is true in our case, see expression (1)), $J_3$ is convex with respect to $(c, \sigma)$ [5]. So $J_3$ has only one global minimum. Huber proposes to alternately minimise $J_3$ with respect to $c$ and $\sigma$. $c$ being fixed, $\sigma$ is estimated using:

$$\left(\sigma^{(m+1)}\right)^2 = \frac{1}{n \cdot a} \sum_{i=1}^{n} \chi(\epsilon_i^{(m)}/\sigma^{(m)}) \left(\sigma^{(m)}\right)^2 \tag{8}$$

where $\chi(x) = x \cdot \rho'(x) - \rho(x)$. The value of $a$ is chosen such that (8) reduces to the standard variance estimate in the quadratic case $\rho(x) = \frac{x^2}{2}$ [5]. Note that, after convergence, the solution of (8) minimises (7) with respect to $\sigma$. Including the scale estimation step, for convex $\rho$ functions, algorithm (5) becomes:

$$\begin{cases} \forall i \in \{1 \ldots n\} \; \epsilon_i^{(m)} = e_i - \sum_{j=1}^{t} c_j^{(m)} \cdot U_{ij} \\ \forall i \in \{1 \ldots n\} \; b_i^{(m+1)} = \frac{\rho'(\epsilon_i^{(m)}/\sigma^{(m)})}{2 \cdot (\epsilon_i^{(m)}/\sigma^{(m)})} \\ (U^T \cdot B^{(m+1)} \cdot U) \cdot c^{(m+1)} = U^T \cdot B^{(m+1)} \cdot e \\ \left(\sigma^{(m+1)}\right)^2 = \frac{1}{n \cdot a} \sum_{i=1}^{n} \chi(\epsilon_i^{(m)}/\sigma^{(m)}) \left(\sigma^{(m)}\right)^2 \end{cases} \tag{9}$$

Huber shows that, when $c$ is fixed, the scale step (8) decreases $J_3$ under particular conditions on $\rho$ ([5]p.180). These conditions are satisfied by the convex function we use in this work. Moreover, when $\sigma^{(m)}$ is fixed, computing $(c^{(m+1)}, b^{(m+1)})$ also decreases $J_3$ [3]. Since the latter is bounded below, algorithm (9) will converge to its (global) minimum. Let us notice that algorithm (9) is only valid for convex functions. In practice, one is also interested in using non-convex functions, which show better outliers rejection capabilities. This point will be discussed in section 5, in which we present a use of M-estimators in continuation.

## 4 Extension to colour image

We now consider the case of colour images. Colour images are transformed into 1-D vectors by concatenating red, green and blue values and a PCA is applied to the $3n \times 3n$ covariance matrix. A first important point for the reconstruction is the definition of residual $\epsilon_i$. Since colour is a discriminant feature for recognition, the red, green and blue components of the pixel must not be considered separately. Therefore, a straightforward extension of (1) is not suitable. Rather, we consider colour components in a pixel-oriented fashion and define the residuals as:

$$\epsilon_i = [(e_i - e_i^*)^2 + (e_{i+n} - e_{i+n}^*)^2 + (e_{i+2n} - e_{i+2n}^*)^2]^{\frac{1}{2}} \tag{10}$$

Approximating their distribution by a gaussian density model, the least square estimation objective function is:

$$J_0 = \sum_{i=1}^{n} \epsilon_i^2 = \sum_{i=1}^{3n} (e_i - e_i^*)^2 \tag{11}$$

It is also possible, as before, to take into account outliers, using the theory previously developed in section 2 and 3, and the resulting algorithm is similar to (9). Although the expression of the residuals is no longer linear with respect to $c_j$, $J_3$ is still convex with respect to $(c, \sigma)$ when is $\rho$ is convex. We sketch the proof in Appendix A.

## 5 M-estimation in continuation

The half-quadratic theory is valid for a large class of functions, defined in [3]. Three standard robust functions are presented in Table 1. Convex functions, like HS, yield a unique solution, but the corresponding *influence functions*, $\rho'$ are monotone. The influence of outliers is then bounded, but not null. From this point of view, hard redescenders [13], as GM, are much more attractive. Unfortunately, hard redescenders yield non-convex energies. Efficient deterministic algorithms can, however, be defined in this case using a gradual approach of non-convexity. We propose here to use (in continuation) the convex function HS with the algorithm (9), using the least square estimate $(c_Q, \sigma_Q)$ as an initial guess, next a non-convex soft redescender HL at a fixed scale using (5), and finally a non-convex hard redescender, GM, at a fixed scale using (5). This strategy is similar to GNC that has already been used in [2][1] for object recognition, with the GM function. However, our algorithm shows two major improvements:

- In GNC, non-convexity is progressively introduced by tuning the scale parameter. At the first step, the energy is made convex on the variation domain of $\epsilon$, by choosing a large value for $\sigma$. This value is related to the largest expected outlier [1], which must be known in advance. In our case, the first function is convex, independently from the values of residuals. Furthermore, the scale parameter is automatically estimated, so no user interaction is required.
- Secondly, exploiting the half-quadratic theory, our algorithm explicitly addresses the problem of the non-quadraticity of the energy. Our approach yields a natural linearization of the normal equations.

The scale parameter $\sigma_{HS}$ is estimated by the algorithm (9). The value of $\sigma_{HL}$ and $\sigma_{GM}$ are computed from $\sigma_{HS}$ in order to match inflexion points of the corresponding $b$ functions.

## 6 Experimental results

As an illustration, we apply here the robust estimation scheme to a road sign recognition problem. A training set of 43 synthetic RGB images of road signs has been collected. Road signs are sampled using a triangular mask, so the background does not affect the estimation. Some training images are shown in figure 1. Only 21 out of 42 eigenvectors are used in this experiment which represent 90% of the information on the database. The recognition is performed using the robust chain. The recognition step consists in finding the training image which is the "closest" to $e^*$. Two control parameters are defined to evaluate the quality of recognition. $d_1$ is the Euclidean distance to the closest training image. Denoting $d_2$, the distance between $e^*$ and the second closest training image, we define a recognition contrast $C$ by $C = \left(1 - \frac{d_1}{d_2}\right) \times 100$.

Figure 2 shows the test images, the results of the four steps of the robust estimation scheme along with the outlier mask (outliers appear in grey) and the recognised model. The first test image $p_1$, was extracted from a real image using a correlation-based algorithm. There is no occlusion in $p_1$, but the road sign is slightly different (in size, orientation, position of the symbol) from the synthetic training models. This explains that the contours appear as outliers. Images $p_2$ to $p_5$ are created by degrading $p_1$ with different synthetic effects. Image $p_2$ presents two occlusions (blue colour) which, although small in size, confuse the least square estimation (the model $m_2$ is recognized instead of $m_1$). This is the same for the tested image $p_3$ whose red inscriptions make the standard estimation identify the model $m_4$. In both cases, the correct model is recognised at the first robust step HS, and the quality of recognition is improved by HL and GM estimations. Images $p_4$ and $p_5$ present the same large occlusion areas with different colours. Note that $p_5$ is a trap, even for human perception. Indeed, black occlusion areas can either be considered as inliers for models $m_5$ or $m_6$, or as outliers for $m_1$. Therefore, the robust chain fails to identify $p_5$. There is not such ambiguity for $p_4$, since blue pixels are clearly identified as outliers. This example shows that colour information is well taken into account by our method.

## 7 Conclusion

We have presented a robust eigenspace recognition method for colour images using M-estimators in continuation. Exploiting the Half-Quadratic theory and results from robust statistics, we propose a non-supervised algorithm which is simple to implement. Our experiments for a colour road sign database show sensible improvements with respect to standard eigenspace approaches.

## A Proof of the convexity of $J_3$ in the RGB case

Following Huber [5]p178, in order to demonstrate that $J_3$ is convex in $(c, \sigma)$, we assume that $(c, \sigma)$ depends linearly on some real parameter $t$ and calculate the second derivative with respect to $t$ of the summand of $J_3$ (omitting subscript $i$). It comes $q(t) = \sigma \cdot \rho(\epsilon/\sigma) + a \cdot \sigma$. By differentiation with respect to $t$, we obtain:

$$\dot{q}(t) = \dot{\sigma}\ \rho(\epsilon/\sigma) + \rho'(\epsilon/\sigma)\left(\dot{\epsilon} - \frac{\epsilon\ \dot{\sigma}}{\sigma}\right) + a\ \dot{\sigma}$$
$$\ddot{q}(t) = \ddot{\epsilon}\ \rho'(\epsilon/\sigma) + \rho''(\epsilon/\sigma)\ \left(\dot{\epsilon} - \frac{\epsilon\ \dot{\sigma}}{\sigma}\right)^2$$

We rewrite $\epsilon$ as:

$$\epsilon = \left\{(e^R - f^R(c))^2 + (e^G - f^G(c))^2 + (e^B - f^B(c))^2\right\}^{\frac{1}{2}}$$

where $f_R$, $f_G$ and $f_B$ are linear in $c$. It comes:

$$\dot{\epsilon} = \frac{-1}{\epsilon}\left\{\dot{f}^R(c)(e^R - f^R(c)) + \dot{f}^G(c)(e^G - f^G(c)) + \dot{f}^B(c)(e^B - f^B(c))\right\}$$

and $\ddot{\epsilon}$ can be expressed as a sum of squares :

$$\begin{aligned}\ddot{\epsilon} = \tfrac{1}{\epsilon^3}\{&\left(\dot{f}^R(c)(e^G - f^G(c)) - \dot{f}^G(c)(e^R - f^R(c))\right)^2\\ &+ \left(\dot{f}^B(c)(e^G - f^G(c)) - \dot{f}^G(c)(e^B - f^B(c))\right)^2\\ &+ \left(\dot{f}^R(c)(e^B - f^B(c)) - \dot{f}^B(c)(e^R - f^R(c))\right)^2\}\end{aligned}$$

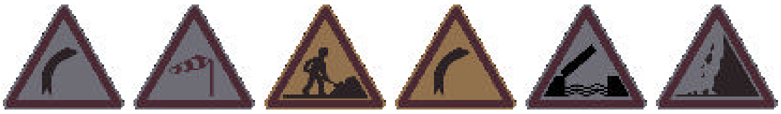
Therefore $\ddot{q} > 0$, and $J_3$ is convex in $(c, \sigma)$. This result holds for a residual expressed with more than 3 components per pixel.

## References


1. M.J. Black and P. Anandan, "A Framework for the robust estimation of optical flow", Proc. 4th ICCV, pp. 231-236, Berlin, May 1993.
2. M. J. Black and A. D. Jepson, "EigenTracking: Robust Matching and Tracking of Articulated Objects Using a View-Based Representation", Int. J. Computer Vision, vol. 26, n1, pp. 63-84,1996.
3. P. Charbonnier, L. Blanc-Féraud, G. Aubert and M. Barlaud, "Deterministic Edge-Preserving Regularization in Computed Imaging", IEEE Transaction Image Processing, vol. 6, n 2, pp. 298-311, February 1997.
4. D. Geman and C. Yang, "Nonlinear image recovery with half-quadratic regularization and FFT's", IEEE Trans. Image Processing, vol. 4, pp.932-946, July 1995.
5. P.J. Huber, "Robust Statistics", John Wiley & Sons, New York, 1981.
6. P. Meer, D. Mintz, A. Rosenfeld and D.Y. Kim, "Robust Regression Methods for Computer Vision: A Review", International Journal of Computer Vision, vol. 6, n 1, pp. 59-70, 1991.
7. B. Moghaddam and A. Pentland, "Probabilistic Visual Learning for Object Recognition", IEEE Transaction on Pattern Analysis an Machine Intelligence, vol. 19, n 7, pp. 696-710, July 1997.
8. H. Murase and S. K. Nayar, "Visual Learning and Recognition of 3D Objects from Appearance", International Journal of Computer Vision, vol.14, pp. 5-24,1995.
9. A. Papoulis, "Probability, Random Variables, and Stochastic Processes", 3rd Edition McGraw-Hill.
10. W. H. Press, W. T. Vetterling, S. A. Teukolsky and B. P. Flannery, "Numerical Recipes in C. The art of scientific computing", 2nd edition Cambridge University Press.
11. P. J. Rousseuw, "Least Median of Squares Regression", Journal of the American Statistical Association, vol. 79, n 388, pp. 871-880, Theory and Methods Section, December 1984.
12. C. V. Stewart, "MINPRAN: A New Robust Estimator for Computer Vision", IEEE Transaction on Pattern Analysis and Machine Intelligence, vol. 17, n 10, pp. 925-938, October 1995.
13. C. V. Stewart, "Bias in Robust Estimation Caused by Discontinuities and Multiple Structures", IEEE Transaction on Pattern Analysis and Machine Intelligence, vol. 19, n 8, pp.818-833, August 1997.

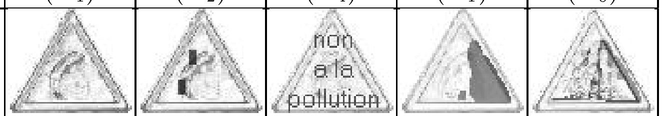

| | $\rho(x)$ | $\rho'(x)$ | convexity |
|---|---|---|---|
| HS | $2\sqrt{1+x^2}-2$ | monotone | convex |
| HL | $\log(1+x^2)$ | soft redescender | non convex |
| GM | $\frac{x^2}{1+x^2}$ | hard redescender | non convex |

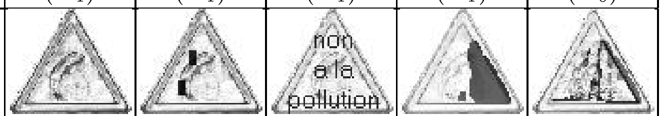
**Table 1.**

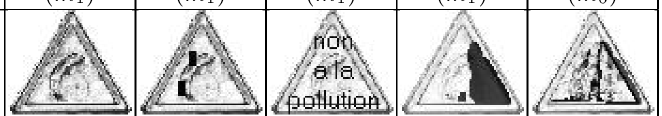


| $m_1$ | $m_2$ | $m_3$ | $m_4$ | $m_5$ | $m_6$ |
|---|---|---|---|---|---|
| (white) | (white) | (yellow) | (yellow) | (white) | (white) |

**Fig. 1.** Sample training images of the road signs database (with indication of their background colour)


| Tested images | $p_1$ | $p_2$ | $p_3$ | $p_4$ | $p_5$ |
|---|---|---|---|---|---|
| Q | $C = 92$ $d_1 = 0.02$ $(m_1)$ | $C = 1.5$ $d_1 = 0.12$ $(m_2)$ | $C = 24$ $d_1 = 0.15$ $(m_4)$ | $C = 36$ $d_1 = 0.21$ $(m_1)$ | $C = 63$ $d_1 = 0.24$ $(m_6)$ |
| HS | $C = 94$ $d_1 = 0.02$ $(m_1)$ | $C = 64$ $d_1 = 0.06$ $(m_1)$ | $C = 47$ $d_1 = 0.12$ $(m_1)$ | $C = 44$ $d_1 = 0.16$ $(m_1)$ | $C = 70$ $d_1 = 0.21$ $(m_6)$ |
| HL | $C = 94$ $d_1 = 0.02$ $(m_1)$ | $C = 81$ $d_1 = 0.04$ $(m_1)$ | $C = 62$ $d_1 = 0.09$ $(m_1)$ | $C = 49$ $d_1 = 0.14$ $(m_1)$ | $C = 75$ $d_1 = 0.19$ $(m_6)$ |
| GM | $C = 94$ $d_1 = 0.02$ $(m_1)$ | $C = 86$ $d_1 = 0.03$ $(m_1)$ | $C = 66$ $d_1 = 0.08$ $(m_1)$ | $C = 53$ $d_1 = 0.13$ $(m_1)$ | $C = 79$ $d_1 = 0.17$ $(m_6)$ |

**Fig. 2.** Results of the robust eigenspace recognition method.