Dense matching using correlation: new measures that are robust near occlusions

Sylvie Chambon and Alain Crouzil
TCI, IRIT, UPS, 118 route de Narbonne
31062 Toulouse Cedex 4, FRANCE
{chambon,crouzil}@irit.fr

Abstract

In the context of computer vision, matching can be done using correlation measures. This paper presents the classification of fifty measures into five families. In addition, eighteen new measures based on robust statistics are presented to deal with the problem of occlusions. An evaluation protocol is proposed and the results show that robust measures (one of the five families), including the new measures, give the best results near occlusions.

1 Introduction

One of the goals of stereovision is to find the third dimension from two images taken from two different angles. While searching for the third dimension, two other problems occur: calibration and matching. Matching is an important task in computer vision, the accuracy of the three-dimensional reconstruction depending on the accuracy of the matching. The problems of matching are: intensity distortions, noises, untextured areas, foreshortening and occlusions. A lot of matching algorithms have been proposed and compared [7, 8, 18] to take these problems into account. The present paper only deals with matching using correlation measures.

In this context, we consider that a correlation measure evaluates the similarity between two data sets: two pixels and their neighbourhoods. Some of the correlation measures (classical measures, derivative-based measures and ordinal measures) have been studied and compared [1, 2, 7, 12, 21, 22]. However, the choice of one correlation measure is difficult so they must first be classified. Here, we are particularly concerned with the occlusion problems and we want to determine the most efficient measures near occlusions.

First, the commonly used correlation measures are presented and classified into five families. The measure properties are also given. Second, eighteen new correlation measures that are robust near occlusions are proposed. In a scene, depth discontinuities induce occlusion problems. Pixels with a different depth from the pixel being studied may be considered as outliers. The tools of robust statistics are insensitive to outliers, so, our measures are based on robust statistics. Third, we set up an evaluation protocol that compares all measures. The results are discussed and conclusions drawn.

2 Taxonomy of the measures

The two grey level images are denoted by I^l and I^r and the following notations are used:

- The size of the correlation windows is: $(2n+1)\times(2m+1)$ and $N=(2n+1)(2m+1)$, $n, m \in \mathbb{N}^*$ and $I_{\max}$ is the maximal grey level;

BMVC 2003 doi:10.5244/C.17.15

- $I^l_{i,j}$ and $I^r_{x,y}$ are the grey levels of the pixels in the left and right images of coordinates (i,j) and (x,y) and $\nabla^v_{x,y}$ are the gradient vectors at a pixel (x,y) in the image $v = l, r$;
- The transposed vector $\mathbf{u}$ is $\mathbf{u}^T$, $\lfloor x \rfloor$ is the integer part of x, $\mathrm{rank}(x)$ is the rank of x, $\mathrm{card}(X)$ is the number of elements in set X and the concatenation is noted: $\otimes$;
- The vectors $\mathbf{f_v}$, $\mathbf{v} = \mathbf{l}, \mathbf{r}$, contain the grey levels of the pixels in the left and right correlation windows: $\mathbf{f_v} = (\cdots I^v_{i+p,j+q} \cdots)^T$, $p \in [-n;n]$, $q \in [-m;m]$;
- If $\sum_{p=-n}^{n} \sum_{q=-m}^{m} = \sum_A$ then the L_P norms are defined by: $\|\mathbf{f_v}\|_P = (\sum_A |I^v_{i+p,j+q}|^P)^{1/P}$ $P \in \mathbb{N}^*$. The Euclidean norm is noted: $\|\mathbf{f_v}\| = \|\mathbf{f_v}\|_2$. The scalar product is defined by: $\mathbf{f_l} \cdot \mathbf{f_r} = \sum_A I^l_{i+p,j+q} I^r_{x+p,y+q}$. The means are noted: $\overline{\mathbf{f_v}} = 1/N \sum_A I^v_{i+p,j+q}$. The variances are defined by: $\mathrm{var}(\mathbf{f_v}) = \overline{(\mathbf{f_v} - \overline{\mathbf{f_v}})^2}$. Element i of vector $\mathbf{f_v}$ is noted f^i_v ;
- The Hamming distance is defined by: $\mathrm{D}_{Ham}(\mathbf{f_l}, \mathbf{f_r}) = \sum_{i=0}^{N-1} \mathrm{sgn}\,|f^i_l - f^i_r|$, with $\mathrm{sgn}(x) = 0$ if $x > 0$, 1 if $x = 0$ or -1 otherwise.

The measures were classified into five families: cross-correlation-based measures, classical statistics-based measures, derivative-based measures, ordinal measures and robust measures. In tables 1, 2, 3, 4 and 7, the invariance properties (column P) are given. Measures without property are noted 0. The properties are with the scalars a, b, c and d:

- $\mathrm{M}(\mathbf{f_l} + a, \mathbf{f_r} + b) = \mathrm{M}(\mathbf{f_l}, \mathbf{f_r})$ (noted 1);
- $\mathrm{M}(a\mathbf{f_l}, b\mathbf{f_r}) = \mathrm{M}(\mathbf{f_l}, \mathbf{f_r})$ (noted 2);
- $\mathrm{M}(a\mathbf{f_l} + b, c\mathbf{f_r} + d) = \mathrm{M}(\mathbf{f_l}, \mathbf{f_r})$ (noted 3).

The abreviations (column ABR.), the measure type (column T): similarity (S) or dissimilarity (D) and the intervals of variation (column I) are also given. These intervals are composed of a lower and an upper bound. In the following description, when no explicit reference is given, the reader should consult [1].

Cross-correlation-based measures Three measures use the scalar product (Table 1).

NAME	ABR.	DEFINITION	I	T	P
Normalised Cross-Correlation	NCC	$(\mathbf{f_l} \cdot \mathbf{f_r}) / \|\mathbf{f_l}\| \|\mathbf{f_r}\|$	$[0;1]$	S	2
Zero mean Normalised Cross-Correlation	ZNCC	$\mathrm{NCC}(\mathbf{f_l} - \overline{\mathbf{f_l}}, \mathbf{f_r} - \overline{\mathbf{f_r}})$	$[-1;1]$	S	3
Moravec	MOR	$2(\mathbf{f_l} - \overline{\mathbf{f_l}}) \cdot (\mathbf{f_r} - \overline{\mathbf{f_r}}) / (\|\mathbf{f_l} - \overline{\mathbf{f_l}}\|^2 + \|\mathbf{f_r} - \overline{\mathbf{f_r}}\|^2)$	$[-1;1]$	S	0

Table 1: Cross-correlation-based measures.

Classical statistics-based measures Eight measures use the distances (can be normalised, centered or locally scaled), the variances [3] and the kurtosis [17] (Table 2).

NAME	ABR.	DEFINITION	I	T	P
Distances	D_P	$\|\mathbf{f_l} - \mathbf{f_r}\|_P{}^P$	$[0; {I_{\max}}^P N]$	D	0
Normalised Distances	ND_P	$\mathrm{D}_P(\mathbf{f_l}, \mathbf{f_r}) / \sqrt{\|\mathbf{f_l}\|_P{}^P \|\mathbf{f_r}\|_P{}^P}$	$[0; {I_{\max}}^P N]$	D	0
Zero mean Distances	ZD_P	$\mathrm{D}_P(\mathbf{f_l} - \overline{\mathbf{f_l}}, \mathbf{f_r} - \overline{\mathbf{f_r}})$	$[0; {I_{\max}}^P N]$	D	1
Zero mean Normalised Distances	ZND_P	$\mathrm{ND}_P(\mathbf{f_l} - \overline{\mathbf{f_l}}, \mathbf{f_r} - \overline{\mathbf{f_r}})$	$[0; {I_{\max}}^P N]$	D	1
Locally Scaled Distances	LSD_P	$\|\mathbf{f_l} - (\overline{\mathbf{f_l}} / \overline{\mathbf{f_r}}) \mathbf{f_r}\|_P{}^P$	$[0; {I_{\max}}^P N]$	D	0
Variance of Differences	VD	$\mathrm{var}(\mathbf{f_l} - \mathbf{f_r})$	$[0; {I_{\max}}^2]$	D	0
Variance of Absolute P-powered Differences	VAD_P	$\mathrm{var}(\|\mathbf{f_l} - \mathbf{f_r}\|^P)$	$[0; {I_{\max}}^{2P}]$	D	0
Kurtosis	K_4	$\|\overline{(\mathbf{f_l} - \mathbf{f_r})^4} - 3\overline{(\mathbf{f_l} - \mathbf{f_r})^2}^2\|$	$[0; {I_{\max}}^4]$	D	0

Table 2: Classical statistics-based measures.

Derivative-based measures Eight measures (Table 3) use the following filters to compute the image derivatives:

- The Pratt [13], Shen-Castan [4], Sobel, Kirsh and Laplacian of Gaussian filters;
- An extension of the Roberts filter:
  $R_r(I_{i,j}) = |I_{i+1,j} - I_{i-1,j}| + |I_{i,j+1} - I_{i,j-1}| + |I_{i+1,j-1} - I_{i-1,j+1}| + |I_{i+1,j+1} - I_{i-1,j-1}|$
  The filtered image is binarised using an adaptive threshold such that 15% of the region of interest (section 3) are greater than this threshold;
- Orientation Code Matching, OCM [19]: $c_{i,j} = \begin{cases} \lfloor \frac{\theta_{i,j}}{\Delta_\theta} \rfloor & \|\nabla I_{i,j}\| > \Gamma \\ L & \text{otherwise} \end{cases}$ with $\theta_{i,j}$ the gradient vector orientation at (i,j) and $c_{i,j}$ lies between 1 and $N' = \frac{2\pi}{\Delta_\theta}$ (N', the number of the levels with a constant width Δ_θ). The term Γ is a pre-specified threshold level to ignore the low contrast pixels and L is a large value assigned as a code for them. In our tests, the constants are: $\Delta_\theta = \frac{\pi}{8}$, $\Gamma = 10$ and $L = 255$ and the distance used is: $D_{ocm}(\mathbf{f_l}, \mathbf{f_r}) = \sum_{i=0}^{N-1} d(f_l^i, f_r^i)$ with $d(a,b) = \begin{cases} \min\{|a-b|, N' - |a-b|\} & \text{if } |a-b| < N' \\ N'/2 & \text{otherwise.} \end{cases}$

The vectors $R_s(\mathbf{f_v})$, $R_k(\mathbf{f_v})$, $R_l(\mathbf{f_v})$, $R_p(\mathbf{f_v})$, $R_r(\mathbf{f_v})$ and $R_{ocm}(\mathbf{f_v})$ are obtained after using the Sobel, Kirsch, Laplacian of Gaussian , Pratt, Roberts filters or the OCM respectively.

NAME	ABR.	DEFINITION	I	T	P
Seitz 1	SES_P	$\|R_s(\mathbf{f_l}) - R_s(\mathbf{f_r})\|_p{}^P$	$[0; {I_{\max}}^P N]$	D	3
Seitz 2	SEK_P	$\|R_k(\mathbf{f_l}) - R_k(\mathbf{f_r})\|_p{}^P$	$[0; {I_{\max}}^P N]$	D	0
Nishihara	NIS	$R_l(\mathbf{f_l}) \cdot R_l(\mathbf{f_r})$	$[0; N]$	S	0
Nack 1	NA_1	$R_r(\mathbf{f_l}) \cdot R_r(\mathbf{f_r}) / N\overline{R_r(\mathbf{f_r})}$	$[0;1]$	S	0
Nack 2	NA_2	$NA_1(\mathbf{f_l}, \mathbf{f_r}) / (N\overline{R_r(\mathbf{f_l})} - R_r(\mathbf{f_l}) \cdot R_r(\mathbf{f_r}))$	$[0;1]$	S	0
Pratt	PRATT	$ZNCC(R_p(\mathbf{f_l}), R_p(\mathbf{f_r}))$	$[-1;1]$	S	0
Orientation code matching	OCM	$(1/N)\, D_{ocm}(R_{ocm}(\mathbf{f_l}), R_{ocm}(\mathbf{f_r}))$	$[0; \frac{N'}{2}]$	S	0
Gradient field correlation	GC	$1 - 2\sum_A \|\nabla^l_{i+p,j+q} - \nabla^r_{k+p,l+q}\| / \sum_A (\|\nabla^l_{i+p,j+q}\| + \|\nabla^r_{k+p,l+q}\|)$	$]-\infty;1]$	S	1

Table 3: Derivative-based measures.

Ordinal measures Six measures (Table 4) use ordered grey levels of the pixels of the correlation window. This family contains three subfamilies:

- The Kaneko measures [10, 11]: $\mathbf{b_v} = (\cdots b_v^i \cdots)^T$ and $\mathbf{c} = (\cdots c^i \cdots)^T$, $i = 0 \cdots N-2$ with $b_v^i = \begin{cases} 1 & \text{if } f_v^{i+1} \geq f_v^i \\ 0 & \text{otherwise} \end{cases}$ and $c^i = \begin{cases} 1 - \|b_l^i - b_r^i\| & \text{if } i = 0 \text{ or } i \text{ even} \\ c^{i-1} & \text{otherwise} \end{cases}$.
  The matrix $\mathbf{C}$ has the weights c^i on its diagonal and 0 elsewhere;
- The Zabih measures [21]: $R_{rank}(\mathbf{f_v}) = \text{card}(\{ f_v^i \mid f_v^i < f_v^{N/2},\ i \in [0; N-1] \})$ and $R_\tau(\mathbf{f_v}) = \bigotimes_i \xi(f_v^{N/2}, f_v^i)$ with $\xi(x,y) = 1$ if $y < x$ or 0 otherwise.
- The Bhat measures [2]: π_v is a permutation with $\pi_v^i = \text{rank}(f_v^i)$, $i \in [0; N-1]$ and $v = l, r$. A permutation composition s is defined by $s^i = \pi_r^k$, $k = (\pi_l^{-1})^i$, where π_l^{-1} is the inverse of π_l. This inverse is defined by: if $\pi_l^i = j$ then $(\pi_l^{-1})^j = i$. The deviation d_m^i for s^i is: $d_m^i = \sum_{j=0}^{i} J(s^j > i)$ where $J(B) = 1$ if B is true or 0 otherwise.

NAME	ABR.	DEFINITION	I	T	P
Increment Sign Correlation	ISC	$(1/(N-1))(\mathbf{b_l} \cdot \mathbf{b_r} + (1-\mathbf{b_l}) \cdot (1-\mathbf{b_r}))$	$[0;1]$	S	3
Selective Coefficient Correlation	SCC	$\mathbf{C}(\mathbf{f_l} - \overline{\mathbf{f_l}}) \cdot (\mathbf{f_r} - \overline{\mathbf{f_r}}) / \|\mathbf{C}(\mathbf{f_l} - \overline{\mathbf{f_l}})\| \|\mathbf{C}(\mathbf{f_r} - \overline{\mathbf{f_r}})\|$	$[0;1]$	S	3
Zabih (L_P norm)	$RANK_P$	$\|R_{rank}(\mathbf{f_l}) - R_{rank}(\mathbf{f_r})\|_p{}^P$	$[0; N^{P+1}]$	D	3
Zabih (Hamming)	CENSUS	$D_{Ham}(R_\tau(\mathbf{f_l}), R_\tau(\mathbf{f_r}))$	$[0; N]$	S	3
Bhat and Nayar 1	κ	$1 - (\max_{i=0 \cdots N-1} d_m^i) / \lfloor N/2 \rfloor$	$[-1;1]$	S	3
Bhat and Nayar 2	χ	$1 - (2d_m^{N/2}) / \lfloor N/2 \rfloor$	$[-1;1]$	S	3

Table 4: Ordinal measures.

Robust measures Twenty four measures (Table 7) use the tools from robust statistics. Among these measures, there are:

- Four partial correlation measures:
  - The Zoghlami measures [23]: The matrices $\mathbf{A_v}$ have, on their diagonals, the weights w_v^i, applied to f_v^i and 0 elsewhere. The transformation used is: $\mathrm{R}_\alpha(\mathbf{f_v}) = \mathbf{A_v}\mathbf{f_v}$ and $\overline{\mathrm{R}_\alpha(\mathbf{f_v})} = 1/N_v \sum_{i=0}^{N-1} \mathrm{R}_\alpha^i(f_v)$ with $N_v = \sum_{i=0}^{N-1} w_v^i$ and $\alpha = F_1, F_2$. ZNCC is used to obtain a map of the maximal scores. With a threshold, the binarisation of this map gives an occlusion map and finally, ZNCC is used again, with the weights of the occlusion map (F_1). The threshold is the mean of the image grey levels. The authors proposed other weights: $1/2(\mathrm{ZNCC}(\mathbf{f_l}, \mathbf{f_r}) + 1)$ (F_2).
  - The Lan measures [12]: The matrices $\mathbf{A}_\alpha$, $\alpha = LMS, MVE$, have on their diagonals, the weights w_α^i applied to f_v^i and 0 elsewhere. The transformation used is: $\mathrm{R}_\alpha(\mathbf{f_v}) = \mathbf{A}_\alpha\mathbf{f_v}$ and $\overline{\mathrm{R}_\alpha(\mathbf{f_v})} = \frac{1}{N_\alpha}\sum_{i=0}^{N-1} \mathrm{R}_\alpha^i(f_v)$ with $N_\alpha = \sum_{i=0}^{N-1} w_\alpha^i$. They use a robust line fitting of N data of two dimensions ($X = \{(f_l^i \ f_r^i)^T\}_{i=0..N-1}$) and the least median of squares (LMS, the smallest median of the squared residuals) or the Minimum Volume Ellipsoid (MVE, the ellipsoid with the smallest volume that contains $h = \lfloor N/2 \rfloor + 1$ points of X);
- The quadrant correlation [9]: $\mathrm{R}_q(\mathbf{f_v}) = \mathrm{sgn}((\mathbf{f_v} - \mathrm{med}(\mathbf{f_v}))/\mathrm{med}\,|\mathbf{f_v} - \mathrm{med}(\mathbf{f_v})|)$;
- The pseudo-norms [6]: the distances (Table 2) with $0 < P < 1$.

First, we suggest to normalise and/or center, like for distances (Table 1), the pseudo-norms (Table 7) in order to make them more invariant. Second, we propose eighteen new measures using the following tools from robust statistics (Table 7):

- The median absolute deviation, *MAD*;
- The least median of squares, the least trimmed squares [15] and the smooth median absolute deviation [14]. We replace the squared differences by the P-powered absolute differences. The ordered values of $\mathbf{x}$ are noted: $(x)_{1:N-1} \leq \ldots \leq (x)_{N-1:N-1}$;
- Six R-estimators [15, 20] (Table 5): $\mathrm{R}_k = \sum_{i=0}^{N-1} a_k(\mathrm{rank}(f_l^i - f_r^i))(f_l^i - f_r^i)$.
  The function a_k is monotonous with: $a_k(0) \leq \ldots \leq a_k(N-1)$, $\sum_{i=0}^{N-1} a_k(i) = 0$ and $a_k(i) = J_k((i+1)/(N+1))$ with $(i+1)/(N+1) \in [0,1]$ and $\int_0^1 J_k(t)\,dt = 0$.

NAME	FUNCTION	NAME	FUNCTION
Wilcoxon	$J_1(t) = t - \frac{1}{2}$	Optimal B-robust estimator	$J_5(t) = \begin{cases} -1.4634 & \text{if } 0 \leq t \leq 0.39 \\ 1.47\phi^{-1}(t) & \text{if } 0.39 \leq t \leq 0.61 \\ 1.4634 & \text{if } 0.61 \leq t \leq 1 \end{cases}$
Bounded normal	$J_2 = \min(1.4634, \max(\phi^{-1}(t), -1.4634))$		
Median	$J_3(t) = \mathrm{sgn}(t - (1/2))$	Minimax	$J_6(t) = \begin{cases} -1.14 & \text{if } 0 \leq t \leq 0.48 \\ \phi^{-1}(\frac{1}{2} + \frac{t-0.5}{t-0.1}) & \text{if } 0.48 \leq t \leq 0.52 \\ 1.14 & \text{if } 0.52 \leq t \leq 1 \end{cases}$
Van der Waerden	$J_4(t) = \phi^{-1}(t)$		

Table 5: The J-functions: ϕ is the $\mathcal{N}(0,1)$ distribution function.

- Eight M-estimators [15, 16, 22] (Table 6): $\mathrm{M}_k(\mathbf{f_l}, \mathbf{f_r}) = \sum_{i=0}^{N-1} \rho_k(f_l^i - f_r^i)$.

NAME	FUNCTION	NAME	FUNCTION
$L_1 - L_2$	$\rho_1(x) = \sqrt{1+x^2}/2 - 1/2$	Cauchy	$\rho_5(x) = \log(1+x^2)$
Fair	$\rho_2(x) = \lvert x\rvert - \log(1+\lvert x\rvert)$	Welsh	$\rho_6(x) = (1 - \mathrm{e}^{-x^2})$
Tukey	$\rho_3(x) = \begin{cases} (1-(1-x^2)^6) & \text{if } \lvert x\rvert \leq 1 \\ 1 & \text{otherwise} \end{cases}$	Huber	$\rho_7(x) = \begin{cases} x^2/2 & \text{if } \lvert x\rvert \leq 1.35 \\ 1.35(\lvert x\rvert - 0.67) & \text{otherwise} \end{cases}$
Geman-McClure	$\rho_4(x) = (x^2/2)/(1+x^2)$	Rousseeuw	$\rho_8(x) = (e^x - 1)/(e^x + 1)$

Table 6: The ρ-functions.

NAME	ABR.	DEFINITION	I	T	P
Zoghlami and Faugeras 1	$ZNCC_2$	$ZNCC(R_{F1}(\mathbf{f_l}), R_{F1}(\mathbf{f_r}))$	$[-1;1]$	S	3
Zoghlami and Faugeras 2	$ZNCC_3$	$ZNCC(R_{F2}(\mathbf{f_l}), R_{F2}(\mathbf{f_r}))$	$[-1;1]$	S	3
Reweigthed Zero mean Sum of Squared Differences	RZSSD	$ZD_2(R_{LMS}(\mathbf{f_l}), R_{LMS}(\mathbf{f_r}))$	$[0; {I_{\max}}^2 N]$	D	1
Reweigthed Zero mean Normalised Cross-Correlation	RZNCC	$ZNCC(R_{MVE}(\mathbf{f_l}), R_{MVE}(\mathbf{f_r}))$	$[-1;1]$	S	3
Quadrant correlation	QUAD	$ZNCC(R_q(\mathbf{f_l}), R_q(\mathbf{f_r}))$	$[0;1]$	S	3
Pseudo-norm	$PSEUDO_P$	$D_P(\mathbf{f_l}, \mathbf{f_r})$, $0 < P < 1$	$[0; {I_{\max}}^P N]$	D	0
Median Absolute Deviation	MAD	$\text{med}\lvert(\mathbf{f_l} - \mathbf{f_r}) - \text{med}(\mathbf{f_l} - \mathbf{f_r})\rvert$	$[0; I_{\max}]$	D	3
Least Median of Powers	LMP_P	$\text{med}(\lvert\mathbf{f_l} - \mathbf{f_r}\rvert^P)$	$[0; {I_{\max}}^P]$	D	3
Least Trimmed Powers	LTP_P	$\sum_{i=0}^{h-1}(\lvert\mathbf{f_l} - \mathbf{f_r}\rvert^P)_{i:N-1}$	$[0, {I_{\max}}^P h]$	D	3
Smooth Median Powered Deviation	$SMPD_P$	$\sum_{i=0}^{h-1}(\lvert\mathbf{f_l} - \mathbf{f_r} - \text{med}(\mathbf{f_l} - \mathbf{f_r})\rvert^P)_{i:N-1}$	$[0, {I_{\max}}^P h]$	D	3
R-estimators	R_k	$\sum_{i=0}^{N-1} J_k(\text{rank}(\frac{f_l^i - f_r^i}{N-1}))(f_l^i - f_r^i)$	$\mathbb{R}$	D	3
M-estimators	M_k	$\sum_{i=0}^{N-1} \rho_k(f_l^i - f_r^i)$	$[0; +\infty[$	D	0

Table 7: Robust measures.

Properties of the measures The first three families are not robust near occlusions because they use the tools of classical statistics that are very sensitive to outliers. Only centered and/or normalised measures, SES_P and GC support intensity distorsions. As the ordinal measures use ranks, they support intensity distortions. They also tolerate outliers. But, these measures can be ambiguous, for example, with these two vectors: $\mathbf{u} = (0\ 1\ 22\ 35\ 46\ 58\ 61\ 121\ 189)^T$ and $\mathbf{v} = (0\ 2\ 42\ 60\ 81\ 100\ 123\ 124\ 125)^T$. When the two vectors are the same (**u** and **u**) and when they are different (**u** and **v**), the optimal score is obtained: these measures can find erroneous correspondences. The robust measures support intensity distortions. All the new measures are robust near occlusions because they use robust estimators. The ρ-functions of M_k do not have a large value for a large difference of grey level which induces the robustness of M_k. To be robust, the ρ-functions should be constant for large values and grow slowly (like $f(x) = x$) for small values.

3 Evaluation protocol

Five pairs of images with ground truth are used but, for space constraints, only three pairs are presented:

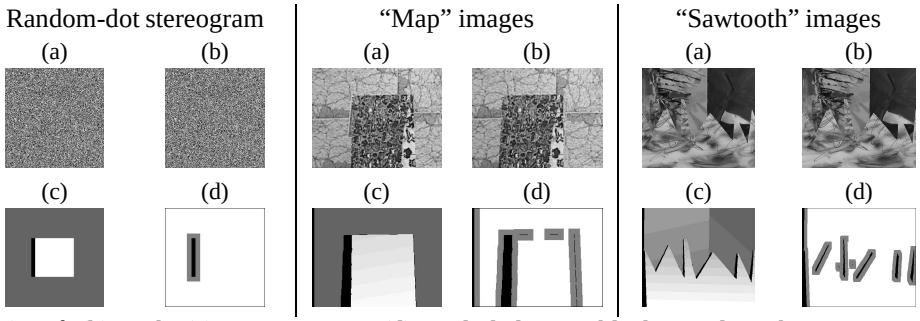
- The random-dot stereogram: These synthetic images (256×256) contain two planes with an occlusion on the left of the closest plane;
- The "map" images (286×216) and the "sawtooth" images (434×380) [18]: These real images were made up piecewise of planar objects (typically posters or paintings, some with cut-out edges).

In the disparity maps, the clearer the pixel is, the closer the point to the image plane and the larger the disparity. The black pixels are occluded pixels. All the pairs of real images can be found at: `http://www.middlebury.edu/stereo/data.html`.

Seventeen criteria were chosen:

- Percentage of correct and false matches;
- Percentage of accepted matches: if the distance between the calculated and the true correspondent is one pixel then the calculated correspondent is accepted. When the percentage of correct matches is low, if this criterion is large then the measure gives a good estimation of the disparities;

- Percentage of false positives and false negatives: the measure finds the pixel is matched whereas it is not matched and vice versa;
- Maximum and mean squared errors (in pixels): maximum or mean Euclidean norms between the calculated matched pixels and the true matched pixels;
- Percentage of correct matched pixels in occluded areas: the morphological dilation of the set of pixels with no corresponding pixels in the other image of the pair is considered. The results in the set of pixels without correspondent and in the set of pixels near the pixels without correspondent are distinguished (Figure 1);
- Maximum and mean ambiguity and maximum inaccuracy [5] (Figure 2);
- Execution time and disparity maps.



(a) Left (b) Right (c) Disparity map (d) Occluded areas, black: pixels without correspondent, grey: region around the black pixels set dilated by the correlation window.

Figure 1: Images, ground truth and occluded areas.

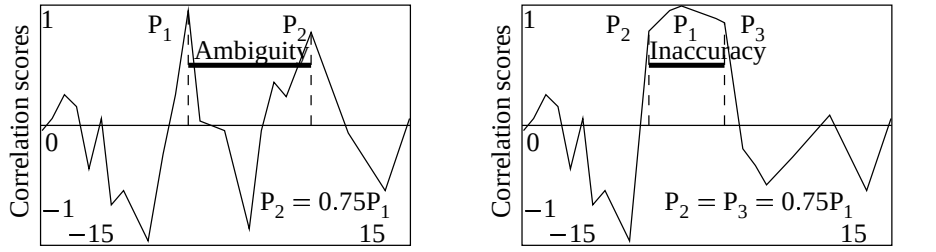


Figure 2: Ambiguity and inaccuracy computations.

Our algorithm is minimal to highlight only the measure behavior. The parameters of the algorithm are the size of the correlation window and the region of interest. A square correlation window (the size grows from 3×3 to 25×25) and a region of interest limited to the size 61×1 (30 pixels before and 30 pixels after the pixel of interest) are chosen. For each pixel in the left image, the algorithm is:

1. The region of interest is determined in the right image;
2. For each pixel in the region of interest, the correlation score is evaluated;
3. The pixel giving the largest score is the matched pixel.

This algorithm only uses similarity measures so each dissimilarity measure is changed into a similarity measure by taking the opposite. Moreover, a bidirectional constraint is added in order to try to locate the occluded pixels. The correlation is performed twice by reversing the roles of two images. The matches for which the reverse correlation falls on the initial point in the left image are considered as valid.

4 Experimental results

In tables 8 to 10, the following abreviations are used: correct matches (COR), false matches (FAL), accepted matches (ACC), false positives (FPOS) , false negatives (FNEG), maximal and mean squared errors (MASE, MSE), maximal and mean ambiguity (MAA, MA), maximal inaccuracy (MAI), correct in the dilated part of the set of occluded pixels (DIL), in the set of occluded pixels (OCC), in the set of pixels near occluded pixels (NOCC) and execution time (TPS). The results are given for the best window size. For each family, one measure that gives the best results is shown (for the robust measures, two measures). In the tables, the best result is emphasised for each column.

Random-dot stereogram The first and second families give good results in non-occluded regions (the best are D_1 and ZNCC). The derivative-based measures lead to the worst results, except GC (Table 8). The robust measures (in particular RZSSD, RZNCC, $PSEUDO_P$, LTP_2, MAD, $SMPD_2$, LMP_2, M_3, M_4, M_6 and M_8) give the best results, the percentage of correct matches is high and the maximum and mean square errors are low (Table 8). For the execution time, among the results shown in the Table 8, MAD and GC are the most expensive. In fact, the measures κ, χ, $PSEUDO_P$, RZSSD and RZNCC are the most expensive. Generally, normalised and/or centered measures have a larger execution time than the others. So, these measures should be used only when there are intensity variations between the two images. The larger the window size is, the lower the ambiguity and inaccuracy (Figure 3). In the first family, the normalised and centered measures (ZNCC and MOR) are less ambiguous and inaccurate than the normalised measure (NCC) because the variation interval of ZNCC and MOR is larger than that of the measure NCC. The derivative-based measures give ambiguous and inaccurate results, except GC. Among the ordinal measures, SCC gives good results (one of the best) whereas the other measures give worse results than the other families. For the robust measures, with small window (smaller than 7×7), the partial measures, MAD, LMP_P and LTP_P have the worst ambiguity and inaccuracy values whereas the $PSEUDO_P$, M_k give the best results (with SCC).

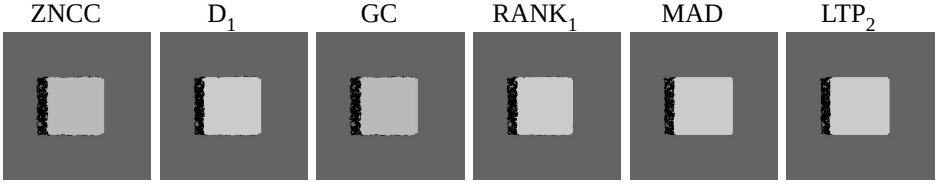


Figure 3: Disparity maps, random-dot stereogram, 7×7.

NAME	COR (%)	FAL (%)	FPOS (%)	FNEG (%)	MASE (pix)	MSE (pix)	MAA (pix)	MA (pix)	MAI (pix)	DIL (%)	OCC (%)	NOCC (%)	TPS (s)
ZNCC	97.6	0.7	**0.2**	1.4	18	0.06	54	**0.11**	**2**	81	**86**	59	**8**
D_1	97.7	0.6	0.4	1.3	**10**	0.06	59	0.14	4	81	78	59	9
GC	96	0.8	0.3	2.9	**10**	0.07	59	0.15	3	83	84	60	40
$RANK_1$	96.6	1	0.3	2.1	**10**	0.09	59	0.15	3	81	79	60	9
MAD	**98.4**	**0.1**	0.3	1.2	**10**	**0**	60	0.16	5	**85**	80	**61**	80
LTP_2	**98.4**	**0.1**	0.4	**1.1**	**10**	**0**	**39**	0.13	8	83	73	**61**	43

Table 8: Random-dot stereogram results, 7×7.

"Map" images The results of the first three families are poor near occluded areas. The ordinal measures lead to good results near occlusions but give the worst disparity maps. The new robust measures provide satisfactory results (Table 9) and are not always more expensive than the others (e.g. $M_{1,3,4,7}$). They have the best results near occlusions (Figure 4, the occluded area in the left of the first plane). The measure LTP_P, having very good results and a clear disparity map, gives the worst results for ambiguity and inaccuracy, with small windows, because it can attribute a high score (close to the maximum score) to two non-correspondent pixels.

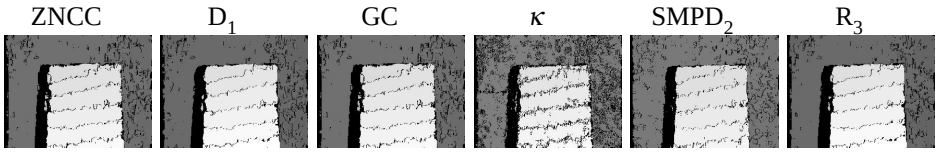


Figure 4: Disparity maps, "map", 9×9.

NAME	COR (%)	ACC (%)	FAL (%)	FPOS (%)	FNEG (%)	MASE (pix)	MSE (pix)	MAA (pix)	MA (pix)	MAI (pix)	DIL (%)	OCC (%)	NOCC (%)	TPS (s)
ZNCC	33	58	58	0.9	8.4	24	0.68	60	0.63	36	67	86	51	**11**
D_1	33	58	59	0.7	7.7	24	0.67	**53**	1.6	46	70	89	54	**11**
GC	33	**59**	60	1	**6.8**	24	0.74	60	**0.38**	**5**	70	82	62	56
κ	32	49	**50**	0.5	19	55	**0.53**	59	0.63	18	70	92	50	419
$SMPD_2$	**35**	56	56	**0.4**	8	**23**	0.6	60	6.02	60	**77**	**93**	**64**	131
R_3	33	58	58	0.7	7.7	24	0.67	54	1.82	46	70	89	54	120

Table 9: "Map" results, 9×9.

"Sawtooth" images Near occlusions, the robust measures are again the most efficient (Figure 5, the occluded area in the left of the "sawtooth"). .

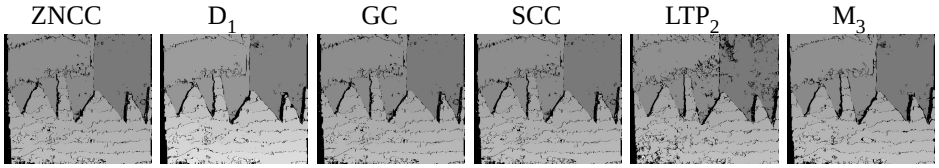


Figure 5: Disparity maps, "sawtooth", 9×9.

NAME	COR (%)	ACC (%)	FAL (%)	FPOS (%)	FNEG (%)	MASE (pix)	MSE (pix)	MAA (pix)	MA (pix)	MAI (pix)	DIL (%)	OCC (%)	NOCC (%)	TPS (s)
ZNCC	53	37	40	1.2	6	38	0.5	**59**	0.56	42	61	69	55	**29**
D_1	54	37	40	1	6	28	0.5	60	2.27	58	64	76	54	**29**
GC	**55**	37	39	1.5	**4**	**26**	0.5	**59**	**0.16**	**11**	62	62	61	155
SCC	53	37	40	1.2	6	38	0.5	60	0.58	41	61	69	55	235
LTP_2	54	35	**38**	**0.6**	8	30	**0.4**	60	6.11	60	**73**	**86**	**64**	225
M_3	54	**38**	39	0.8	6	27	0.5	60	0.82	50	68	81	58	276

Table 10: "Sawtooth" results, 9×9.

Summary of the results Among all the measures studied, those of the two first families, GC (derivative-based measure) and SCC (ordinal measure) give good results. In contrast, derivative-based measures are not efficient. Ordinal measures that are efficient in occluded regions are not really efficient in non-occluded areas. Robust measures are the most efficient particularly partial correlations, $PSEUDO_P$, MAD, LMP_P, LTP_P, $SMPD_P$,

R_k and M_k. However, some of these measures are not as efficient as classical measures in non-occluded areas: the partial correlations, MAD and LMP_P. If the execution time, the ambiguity and the inaccuracy are taken into account, the measures MAD, LMP_P, LTP_P, $SMPD_P$ and R_k are less efficient. Finally, with all the tests that have been done, the M-estimator-based measures lead to the best results, the best disparity maps and a reasonable execution time.

5 Conclusion

Firstly, this work classifies correlation measures into five families. The description of the properties of these measures can help in the choice of a correlation measure. Then, eighteen new robust measures are proposed. The results show the most efficient measures: the robust measures and in particular, all the M-estimator-based measures. Among the proposed measures, some points might be improved: some measures (LTP_P, R_k, MAD and LMP_P) have a high execution time. In fact, the measure implementation was not optimised so the execution times are not the best that can be obtained. A lot of methods can be used to improve this implementation. Moreover, the eighteen new measures can be integrated in a matching algorithm. In fact, robust measures are very efficient near occlusions but some measures, like GC, are more efficient than robust measures in non-occluded areas. So, our future work will be to develop a robust matching algorithm that will use both robust and non-robust measures.

References

[1] P. Aschwanden and W. Guggenbül. Experimental results from a comparative study on correlation type registration algorithms. In Förstner and Ruwiedel, editors, *Robust computer vision: Quality of Vision Algorithms*. Wichmann, Karlsruhe, Germany, March 1992.

[2] D. N. Bhat and S. K. Nayar. Ordinal measures for image correspondence. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(4):415–423, April 1998.

[3] G. S. Cox. Template matching and measures of match in image processing. Technical report, University of Cape Town, South Africa, July 1995. `http://www.dip.ee.uct.ac.za/imageproc/pattern/`.

[4] A. Crouzil, L. Massip-Pailhes, and S. Castan. A new correlation criterion based on gradient fields similarity. In *International Conference on Pattern Recognition*, volume 1, pages 632–636, Vienna, Austria, August 1996.

[5] O. De Joinville, H. Maître, D. Piquet Pellorce, and M. Roux. How to design DEM assessment maps. In *Pattern recognition in Remote Sensing Workshop*, Andorra La Vella, Andorra, September 2000.

[6] J. Delon and B. Rougé. Le phénomène d'adhérence en stéréoscopie dépend du critère de corrélation. In *GRETSI*, Toulouse, France, September 2001. (in French).

[7] A. Giachetti. Matching techniques to compute image motion. *Image and Vision Computing*, 18(3):245–258, February 2000.

[8] L. Gottesfeld Brown. A survey of image registration techniques. *ACM Computing Surveys*, 24(4):325–376, December 1992.

[9] P. J. Huber. *Robust statistics*, chapter 8, pages 204–205. John Wiley & Sons, New-York, USA, 1981.

[10] S. Kaneko, I. Murase, and S. Igarashi. Robust image registration by increment sign correlation. *Pattern Recognition*, 35(10):2223–2234, October 2002.

[11] S. Kaneko, Y. Satoh, and S. Igarashi. Using selective correlation coefficient for robust image registration. *Pattern Recognition*, 36(5):1165–1173, May 2003.

[12] Z. D. Lan and R. Mohr. Robust matching by partial correlation. In *British Machine Vision Conference*, pages 651–660, Birmingham, England, September 1995.

[13] W. K. Pratt. *Digital image processing*, chapter 20, pages 666–667. Wiley-Interscience Publication, 1978.

[14] P. J. Rousseeuw and C. Croux. L_1*-Statistical Analysis and Related Methods*, pages 77–92. Yadolah Dodge, Amsterdam, Holland, 1999.

[15] P. J. Rousseeuw and A. M. Leroy. *Robust regression and outlier detection*. J. Wiley & Sons, New-York, USA, 1987.

[16] P. J. Rousseeuw and S. Verboven. Robust estimation in very small samples. *Computational Statistics and Data Analysis*, 40(4):741–846, October 2002.

[17] M. Rziza, D. Aboutajdine, L. Morin, and A. Tamtaoui. Schéma multirésolution d'estimation d'un champ de disparités dense sous contrainte épipolaire pour les images bruitées. In *GRETSI*, Toulouse, France, September 2001. (in French).

[18] D. Scharstein and R. Szeliski. A taxomomy and evaluation of dense two-frame stereo correspondence algorithms. *International Journal of Computer Vision*, 47(1):7–42, April 2002.

[19] F. Ullah, S. Kaneko, and S. Igarashi. Orientation code matching for robust object search. *IEICE Transactions on Information and Systems*, E-84-D(8):999–1006, March 2001.

[20] Y. Wang and D. Wiens. Optimal, robust R-estimators and test statistics in the linear model. *Statistics and Probability Letters*, 14:179–188, June 1992.

[21] R. Zabih and J. Woodfill. Non-parametric local transforms for computing visual correspondence. In *Proceedings of the European Conference on Computer Vision*, pages 151–158, Stockholm, Sweden, 1994.

[22] Z. Zhang. Parameter estimation techniques: A tutorial with application to conic fitting. *International Journal of Image and Vision Computing*, 15(1):59–76, January 1997.

[23] I. Zoghlami, O. Faugeras, and R. Deriche. Traitement des occlusions pour la modification d'objet plan dans une séquence d'image. In *Actes du congrès francophone de Vision par Ordinateur, ORASIS*, pages 93–103, Clermont-Ferrand, France, May 1996. (in French).