A SUFFICIENT CONDITION BASED ON THE CAUCHY-SCHWARZ INEQUALITY FOR EFFICIENT TEMPLATE MATCHING

Luigi Di Stefano

DEIS-ARCES
University of Bologna (ITALY)
ldistefano@deis.unibo.it

Stefano Mattoccia

DEIS-ARCES
University of Bologna (Italy)
smattoccia@deis.unibo.it

ABSTRACT

The paper proposes a technique aimed at reducing the number of calculations required to carry out an exhaustive template matching process based on the Normalized Cross Correlation (NCC). The technique deploys an effective sufficient condition, relying on the recently introduced concept of bounded partial correlation, that allows rapid elimination of the points that can not provide a better cross-correlation score with respect to the current best candidate. In this paper we devise a novel sufficient condition based on the Cauchy-Schwarz inequality and compare the experimental results with those attained using the standard NCC-based template matching algorithm and the already known sufficient condition based on the Jensen inequality.

1. INTRODUCTION

Template matching is one fundamental task occurring in countless image analysis applications. The basic template matching algorithm consists in calculating at each position of the image under examination a *distorsion* or *correlation* function that measures the degree of similarity between the template and the image. Then, the minimum distorsion, or maximum correlation, position is taken to locate the template into the examined image. Typical distorsion measures are the Sum of Absolute Differences (SAD) and the Sum of Squared Differences (SSD). SAD and SSD are deployed also in applications such as *motion estimation* and *vector quantization*. However, as far as template matching is concerned, Normalized Cross Correlation (NCC) is often the adopted similarity measure due to its better robustness with respect to photometric variations. With large size images and/or templates the matching process can be computationally very expensive and hence numerous techniques aimed at speeding up the basic algorithm have been proposed in literature (see [1] for a concise review).

Among the major ones are 1) the use of multi-resolution schemes: locating a coarse-resolution template into a coarse-resolution image first and then refining the search at the higher resolution levels, 2) sub-sampling the image and the template, 3) two-stage matching: matching a sub-template first, and then the whole template only at good candidate positions, 4) point correlation: matching only a precomputed set of points that are selected optimally in a heuristic way [2].

However, these techniques imply a non exhaustive search process since they do not compare the full resolution image with the full resolution template at every search position and therefore can be trapped by local extremes resulting in wrong localisation of the template. On the other hand, in the case of distorsion measures the SEA (Successive Elimination Algorithm) [3], [4] and PDE (Partial Distorsion Elimination) [5] techniques allow for notably speeding up the computation required by an exhaustive-search template matching process. SEA relies on fast evaluation of a lower-bound for the distorsion measure: if the bounding function exceeds the current minimum, the position can be skipped without calculating the actual distorsion. PDE consists of terminating the evaluation of the distorsion measure if it exceeds the current minimum. As for the Normalized Cross Correlation, we have shown recently that it is possible to speed-up the exhaustive-search template matching process by deploying an upper-bound of the NCC function based on the Jensen inequality [6].

In this paper we propose an improvement to the technique described in [6] that consists in the use of a more effective bounding function derived from the Cauchy-Schwarz inequality.

2. BOUNDING THE NCC

With Normalized Cross Correlation the template sub-image, T, is located into the image, I, under examination by searching for the maximum of the NCC function (1). The numerator of (1) represents the *cross correlation* between the template and the image, $C(x,y)$, and its computation turns out to be the bottleneck in the evaluation of $\eta(x,y)$. In fact, the two terms appearing in the denominator represent the L_2

0-7803-7750-8/03/$17.00 ©2003 IEEE. ICIP 2003

norms of the sub-image under examination, $\|I(x,y)\|_2$, and of the template,$\|T\|_2$. The latter can be computed once at start up, the former can be obtained very efficiently using a box-filtering scheme [7] by accumulating 4 product terms per image point.

$$\eta\left(x,y\right)=\frac{\sum_{j=1}^{N}\sum_{i=1}^{M}I\left(x+i,y+j\right)\cdot T\left(i,j\right)}{\sqrt{\sum_{j=1}^{N}\sum_{i=1}^{M}I\left(x+i,y+j\right)^2}\cdot\sqrt{\sum_{j=1}^{N}\sum_{i=1}^{M}T\left(i,j\right)^2}}$$

Suppose now that a function $\beta(x,y)$ exists such that $\beta(x,y)$ is an upper-bound for $C(x,y)$:

$$\beta(x,y)\geq C(x,y)=\sum_{j=1}^{N}\sum_{i=1}^{M}I\left(x+i,y+j\right)\cdot T\left(i,j\right), \quad (1)$$

by normalising β we obtain an upper-bound for the NCC function:

$$\frac{\beta(x,y)}{\|I(x,y)\|_2\cdot\|T\|_2}\geq\frac{C\left(x,y\right)}{\|I(x,y)\|_2\cdot\|T\|_2}=\eta(x,y) \quad (2)$$

Then, indicating as η_{max} the current correlation maximum, if the following inequality is verified at point x,y

$$\frac{\beta(x,y)}{\|I(x,y)\|_2\cdot\|T\|_2}<\eta_{max} \quad (3)$$

then the matching process can proceed with the next point without carrying out the calculation of $C(x,y)$, for the point is guaranteed not to correspond to the new correlation maximum. Hence, (3) is a sufficient condition for skipping the points that can not improve the degree of matching with respect to the current maximum without calculating the actual cross correlation. Conversely, if (3) is not verified then it is necessary to compute $C(x,y)$, normalise it by the product $\|I(x,y)\|_2\cdot\|T\|_2$ and check the *new maximum condition*:

$$\frac{C(x,y)}{\|I(x,y)\|_2\cdot\|T\|_2}\geq\eta_{max} \quad (4)$$

2.1. The upper bound based on Jensen's inequality

In [6] $\beta(x,y)$ was obtained by exploiting the relationship between geometric and arithmetic series known as Jensen's inequality [8]:

$$\sqrt[n]{\prod_{k=1}^{n}a_k}\leq\frac{\sum_{k=1}^{n}a_k}{n}, \quad (5)$$

where a is n-dimensional vector.

From (5), by means of simple manipulations, it is possible to show that function (6) satisfies (1) at each image point.

$$\beta_J(x,y)=\frac{\sum_{j=1}^{N}\sum_{i=1}^{M}I\left(x+i,y+j\right)^2+\sum_{j=1}^{N}\sum_{i=1}^{M}T\left(i,j\right)^2}{2} \quad (6)$$

2.2. A novel bound based on Cauchy-Schwarz's inequality

Cauchy-Schwarz's inequality provides a relationship between the inner product of two vectors and their L_2 norms, and can be used to devise a more effective bound of the cross correlation term, $C(x,y)$. Given two n-dimensional vectors, a and b, the Cauchy-Schwarz inequality is given by:

$$\sum_{k=1}^{n}a_k\cdot b_k\leq\sqrt{\sum_{k=1}^{n}a_k^2}\cdot\sqrt{\sum_{k=1}^{n}b_k^2} \quad (7)$$

Hence, applying (7) to the terms appearing at the numerator of (1) we can derive a novel function, $\beta_{CS}(x,y)$, that satisfies (1) at each image point:

$$\beta_{CS}(x,y)=\sqrt{\sum_{j=1}^{N}\sum_{i=1}^{M}I\left(x+i,y+j\right)^2}\cdot\sqrt{\sum_{j=1}^{N}\sum_{i=1}^{M}T\left(i,j\right)^2} \quad (8)$$

Though not shown in this paper for the sake of brevity, it can be demonstrated that $\beta_{CS}(x,y)$ upper bounds $C(x,y)$ more tightly than $\beta_J(x,y)$.

Unfortunately, and as it is also the case of $\beta_J(x,y)$, if we simply plug $\beta_{CS}(x,y)$ into (3) we do not get a useful sufficient condition since (3) turns out to be always false. Nevertheless, an effective sufficient condition can be obtained by computing only a small portion of the actual correlation function, referred to as "partial correlation", and bounding the residual portion of the correlation product with a term derived from (8). Hence, defining the partial correlation associated with row n as

$$C(x,y)|_1^n=\sum_{j=1}^{n}\sum_{i=1}^{M}I\left(x+i,y+j\right)\cdot T\left(i,j\right), \quad (9)$$

and the term bounding the residual correlation product as

$$\beta_{CS}\left(x,y\right)|_{n+1}^{N}= \quad (10)$$
$$\sqrt{\sum_{j=n+1}^{N}\sum_{i=1}^{M}I\left(x+i,y+j\right)^2}\cdot\sqrt{\sum_{j=n+1}^{N}\sum_{i=1}^{M}T\left(i,j\right)^2}$$

we obtain the following upper-bound for $C(x,y)$:

$$\gamma(x,y)|_n=C(x,y)|_1^n+\beta_{CS}(x,y)|_{n+1}^{N} \quad (11)$$

Hence, function $\gamma(x,y)|_n$ combines the evaluation of a given part of the total correlation, $C(x,y)|_1^n$, with the approximation of the residual part by means of an upper-bounding term, $\beta_{CS}(x,y)\,|_{n+1}^N$, derived from (8).

Finally, the sufficient condition for skipping the points that cannot improve the correlation with respect to the current maximum is obtained by using $\gamma(x,y)|_n$ in place of $\beta(x,y)$ in (3):

$$\frac{\gamma(x,y)|_n}{\|I(x,y)\|_2 \cdot \|T\|_2} < \eta_m. \tag{12}$$

Evaluating (12) instead of calculating $C(x,y)$ holds the potential for speeding up the matching process as long as (12) can be computed much more rapidly than $C(x,y)$. This can be achieved running a distinct box-filter on each of the two sub-windows shown in Figure 1, with the first box filter calculating the quantity $q_1 = \sum_{j=1}^{n}\sum_{i=1}^{M} I\left(x+i,y+j\right)^2$ and the second the quantity $q_2 = \sum_{j=n+1}^{N}\sum_{i=1}^{M} I\left(x+i,y+j\right)^2$. Then, q_2 can be used to compute the bounding term appearing in (11) (i.e. $\beta_{CS}(x,y)\,|_{n+1}^N$), while q_1 and q_2 can be added together to obtain $\|I(x,y)\|_2$, which is needed to normalise $\gamma(x,y)|_n$ in the sufficient condition (12) and $C(x,y)$ in the new maximum condition (4). The computational cost associated with the described double box-filtering scheme amounts to the accumulation of 8 product terms per point. The evaluation of the partial correlation term present in $\gamma(x,y)|_n$ does not introduce any overhead with respect to the standard algorithm since when (12) fails $C(x,y)$ can be attained adding the residual product term $C(x,y)|_{n+1}^N$ to $C(x,y)|_1^n$. Eventually, the last quantity needed to obtain $\beta_{CS}(x,y)\,|_{n+1}^N$, i.e. $\sum_{j=n+1}^{N}\sum_{i=1}^{M} T\left(i,j\right)^2$, can be precomputed at initialisation time.

We call $C_r = \frac{n}{N}$ (Correlation Ratio) the ratio between the number of points involved in the calculation of the partial correlation and those needed to evaluate the entire correlation function It is clear from the definition of $\gamma(x,y)|_n$ that, as C_r increases, the bounding function gets closer to the actual correlation and the sufficient condition becomes more effective. However, the computational savings associated with skipping points satisfying the sufficient condition amounts basically to the residual fraction of the correlation, which decreases with C_r. Hence, C_r turns out to be the fundamental parameter of the proposed algorithm, its choice implying a trade-off between two opposite requirements: it should be kept large to effectively skip non-maximum points and small to gain a substantial computational benefit at skipped points.

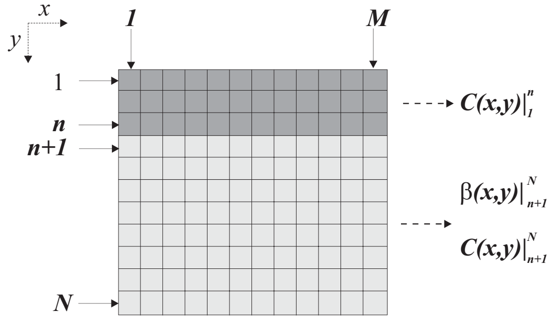

Fig. 1. Partitioning of the template and of the window under examination.

3. EXPERIMENTAL RESULTS

The proposed algorithm based on the Cachy-Schwarz inequality (here referred to as BPC_{CS}), as well as that based on the Jensen inequality [6] (here referred to as BPC_J), belongs to the class of exhaustive search algorithms. Hence, in this section we evaluate the performance of BPC_{CS} and BPC_J, by comparing the number of required operations to those needed by the standard NCC algorithm (here referred to as NCC). We will use as performance evaluation metric, A_{ops}, the ratio between the average number of elementary operations (i.e. accumulation of one product term) per point executed by BPC_{CS}, BPC_J and that required by the standard NCC algorithm.

The number of operations executed by BPC_{CS} and BPC_J at a given point depends on the outcome of their respective sufficient conditions. If this is satisfied, the algorithm executes only the $M \cdot n$ operations needed to evaluate the partial correlation term. If the condition fails, the partial correlation term must be integrated with its residual part, giving a total number of operations equal to $M \cdot N$. In both cases we must add the overhead due to the double box-filtering scheme, which requires 8 additional operations per point. Hence, if $W \times H$ is the image size and S the number of points for which the sufficient condition is satisfied, we can express the total number of elementary operations executed by BPC_{CS} and BPC_J as $T_{ops} = (W \cdot H - S) \cdot M \cdot N + S \cdot M \cdot n + 8 \cdot W \cdot H$.

Experimental results are provided on the data set[1] (i.e. images *pcb1*, *pcb2*, *albert*, *pcb3*,*plants* and *pm*) also considered in [6]. The sub-image used as template has been extracted from a different - though similar - image. Within the entire data set we use $C_r = 30\%$ and four correlation thresholds: 0.0 (i.e. no threshold), 0.95, 0.97 and 0.98. The correlation threshold is a parameter typically used within a

[1]The images can be seen at:
www.vision.deis.unibo.it/~smattoccia/PatternMatching.html

Algorithm	NCC	BPC_J	BPC_{CS}	BPC_J	BPC_{CS}	BPC_J	BPC_{CS}	BPC_J	BPC_{CS}
Threshold	–	**0.00**	**0.00**	**0.95**	**0.95**	**0.97**	**0.97**	**0.98**	**0.98**
pcb1	100.00	68.20	64.90	45.06	33.58	39.82	30.85	38.72	30.84
pcb2	100.00	62.59	43.20	62.58	43.19	54.72	33.89	48.96	31.20
albert	100.00	53.88	36.69	53.17	35.92	51.29	30.65	50.34	30.26
pcb3	100.00	60.42	51.36	60.42	51.36	60.36	50.02	56.94	43.85
plants	100.00	59.73	46.16	49.05	32.70	45.16	30.08	44.16	30.07
pm	100.00	67.88	46.68	60.12	39.04	53.17	31.51	50.76	30.24
Mean	100.00	62.12	48.17	55.07	39.30	50.75	34.50	48.31	32.74

Table 1. Average number of elementary operations per point (A_{ops}) for NCC, BPC_J and BPC_{CS} algorithms. For BPC_J and BPC_{CS} parameter C_r = 30%. The last row shows the mean values of A_{ops} computed within the data set.

template matching algorithm that represents the minimum correlation value that must be found to accept a match.

The Table 1 reports the average number of elementary operation per point executed by standard NCC algorithm, BPC_J and BPC_{CS}. Table 1 shows that both BPC_J and BPC_{CS} yield significant computational savings throughout the considered dataset. Moreover, BPC_{CS} always outperforms BPC_J significantly by virtue of its capability of upper-bounding the cross correlation more tightly. Compared to the standard NCC standard algorithm, BPC_{CS} allows for reducing A_{ops} in the worst case (*pcb1*, threshold 0.00) up to 64.90% and in the best cases (*pcb1*, *pcb2*, *albert*, *plants* and *pm* with threshold 0.98) up to $\simeq 30\%$.

The Table shows also that as the correlation threshold gets more selective the performance provided by BPC_{CS} and BPC_J gets better. In addition, it is worth observing that BPC_{CS} is significantly more effective than BPC_J in taking advantage of a more selective threshold. Finally, the mean values of each column, shown in the bottom row of the Table provide a measure of the global performance of the considered techniques on the addressed dataset. It can be noticed that, in average terms, BPC_{CS} in the worst case halves the number of operation and in the best case (with threshold 0.98) allows for reducing it up to 32.74%. Best and worst results for BPC_J are respectively 62.12% and 48.31%.

4. CONCLUSION

We have shown an algorithm for efficient exhaustive template matching based on the normalized cross correlation. The algorithm deploys a sufficient condition which allows for discarding rapidly the search positions that are guaranteed not to provide a better degree of match with respect to the current best-matching one. The proposed sufficient condition relies on the Cauchy-Schwarz inequality and on the recently introduced concept of bounded partial correlation. The condition can be evaluated very efficiently using a suitable computation scheme requiring a limited number of operations. The performance of the proposed algorithm is compared to the standard NCC algorithm and to another similar algorithm based on the Jensen inequality. Experimental results show that the proposed algorithms clearly outperforms the algorithm based on the Jensen inequality and allows for a significant reduction of the calculations required to carry out exhaustive template matching based on the normalized cross correlation.

5. REFERENCES

[1] A Rosenfeld and A.C. Kak, *Digital Picture Processing*, vol. 2, Academic Press, 1982.

[2] W Krattenthaler, K.J. Mayer, and M Zeiler, “Point correlation: a reduced-cost template matching technique,” in *Proc. of the 1st IEEE International Conference on Image Processing*, Austin, Texas, USA, 1994, vol. 1, pp. 208–212.

[3] W Li and E Salari, “Successive elimination algorithm for motion estimation,” *IEEE Trans. on Image Processing*, vol. 4, no. 1, pp. 105–107, 1995.

[4] H.S. Wang and R.M. Mersereau, “Fast algorithms for the estimation of motion vectors,” *IEEE Trans. on Image Processing*, vol. 8, no. 3, pp. 435–439, 1999.

[5] C.D. Bei and R.M. Gray, “An improvement of the minimum distorsion encoding algorithm for vector quantization,” *IEEE Trans. on Communication*, vol. 33, pp. 1132–1133, 1985.

[6] L Di Stefano and S Mattoccia, “Fast Template Matching using Bounded Partial Correlation,” *Machine Vision and Applications*, vol. 13, no. 4, pp. 213–221, Feb 2003.

[7] M Mc Donnel, “Box-filtering techniques,” *Computer Graphics and Image Processing*, vol. 17, pp. 65–70, 1981.

[8] W Rudin, *Real and Complex Analysis*, McGraw-Hill, New York, USA, 1966.