Grayscale Template-Matching Invariant to Rotation, Scale, Translation, Brightness and Contrast

Hae Yong Kim and Sidnei Alves de Araújo

Escola Politécnica, Universidade de São Paulo, Brazil.

1. Introduction

Goal: Find a grayscale query image *Q* in another image to analyze *A*, invariant to RSTBC:

- **Rotation**
- **Scaling**
- **Translation**
- **Brightness**
- **Contrast**

The proposed solution: Ciratefi (Circular, Radial and Template matching Filter). It consists of three cascaded filters. Each filter excludes pixels that have no chance of matching the template.

Ciratefi makes the classic and well-known template matching to become invariant to scaling and rotation.

Examples:

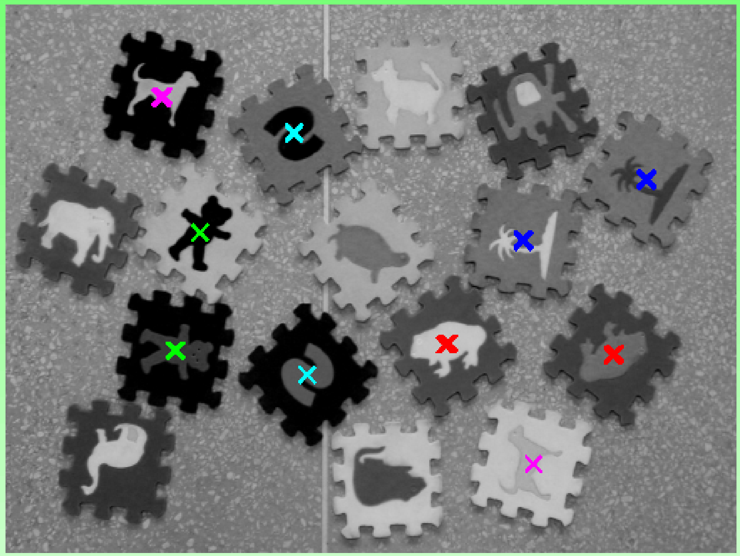



Image © 2007 DigitalGlobe
Google
Pointer 23°34'07.25" S 46°24'55.43" W
Streaming 100%
Eye alt 2008 ft

2.1 Brightness/contrast-invariant template matching

Template matching uses some difference measure to evaluate how well template *Q* matches a position of image *A*.

We use the correlation coefficient r_{xy}, because it is invariant to brightness and contrast.

There is a matching if $r_{xy} > t_B$ (or $|r_{xy}| > t_B$, to allow matching negative template instances).

We assume that the correlation is zero if the brightness correction γ or the contrast correction β is above some threshold.

Brightness/contrast-aware correlation:

$$\mathrm{Corr}(\boldsymbol{x}, \boldsymbol{y}) = \begin{cases} 0, \ \text{if} \ |\beta| \le t_\beta, 1/t_\beta \le |\beta| \ \text{or} \ |\gamma| > t_\gamma \\ r_{\boldsymbol{xy}}, \ \text{otherwise} \end{cases}$$

2.2 Brute force RSTBC-invariant template matching

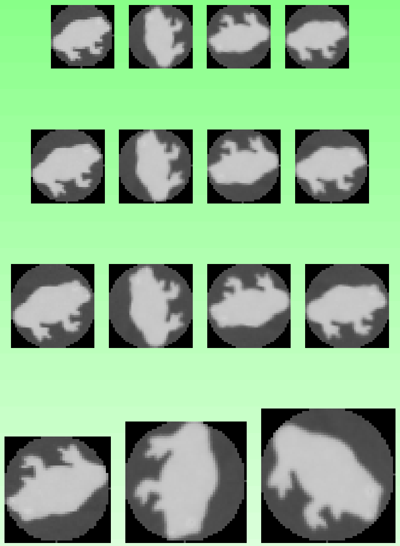
To achieve RSTBC-invariance, template *Q* must be rotated and scaled by all angles and factors and matched against all pixels of *A*.

In practice, discrete sets of angles and scale factors are used. For example:

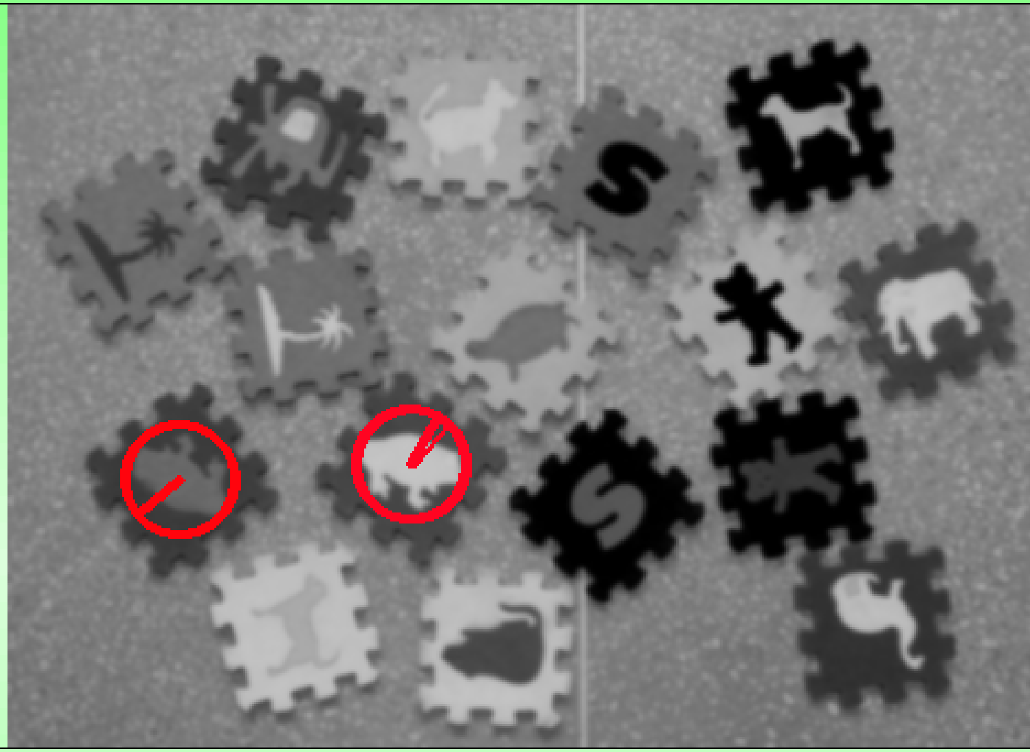
- **(α_0=0, α_1=10, ..., α_{35}=350)**
- **(s_0=0.6, s_1=0.7, ..., s_5=1.1)**

Problem: Brute force template matching is too slow. Brute force takes 2 hours to solve the example below, while Ciratefi takes only 8 seconds.

Note: Images *A* and *Q* are low-pass filtered to avoid that a small misalignment may cause a large mismatching.



Some of the scaled and rotated query images.



RSTBC-invariant brute-force template matching takes 2 hours, while Ciratefi takes only 8 seconds. Circles indicate scales and “watch hands” indicate orientations.

3. Circular sampling filter (Cifi)

Ciratefi consists of 3 cascaded filters: Cifi, Rafi and Tefi.

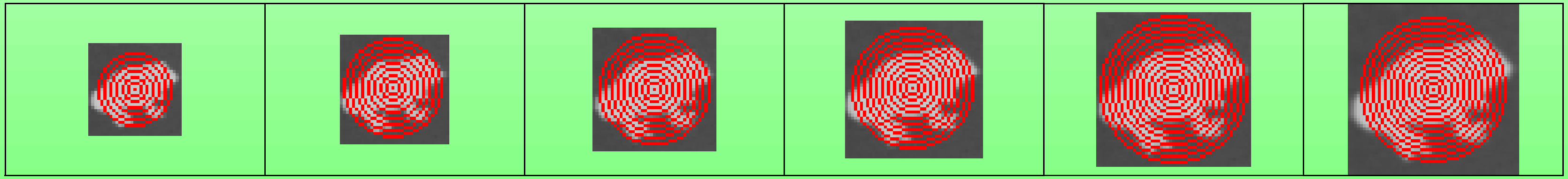
Cifi uses the projections of *A* and *Q* on a set of rings to detect the pixels that have chance of matching the template, named "first grade candidate pixels" and their scale factors.

Circular sampling *Cis* is the average grayscale of the pixels of an image *B* situated at distance *r* from pixel (x,y):

$$\mathrm{Cis}_B(x,y,r)=\int_0^{2\pi} B(x+r\cos\theta,\, y+r\sin\theta)\,d\theta$$

C_Q is a matrix of multi-scale, rotation-invariant features of *Q*:

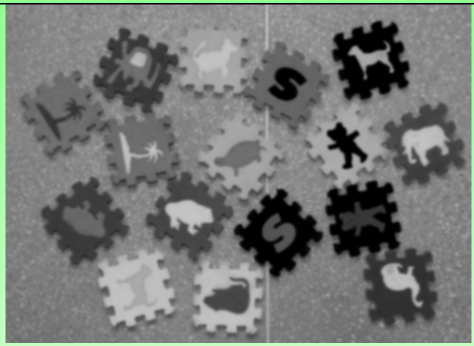
$$C_Q[i,k]=\mathrm{Cis}_{Q_i}(x_0,y_0,r_k),\quad 0\le i<n \text{ and } 0\le k<l$$



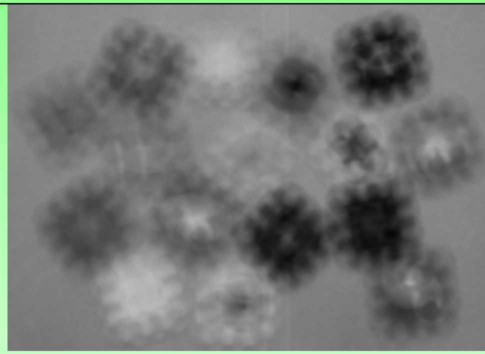
The average grayscales on red circles form the matrix C_Q.

C_A is multi-scale 3-D image with rotation-invariant features of *A*:

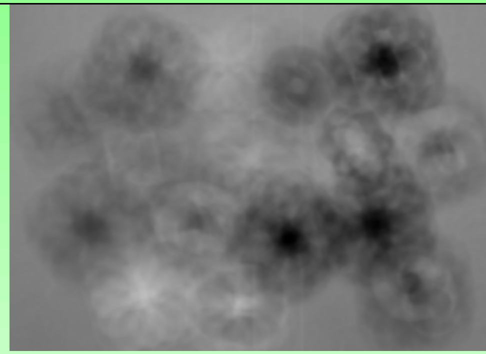
$$C_A[x, y, k] = \mathrm{Cis}_A(x, y, r_k), \quad 0 \le k < l \text{ and } (x, y) \in \mathrm{domain}(A)$$



radius r_k=0 pixels



radius r_k=6 pixels



radius r_k=12 pixels

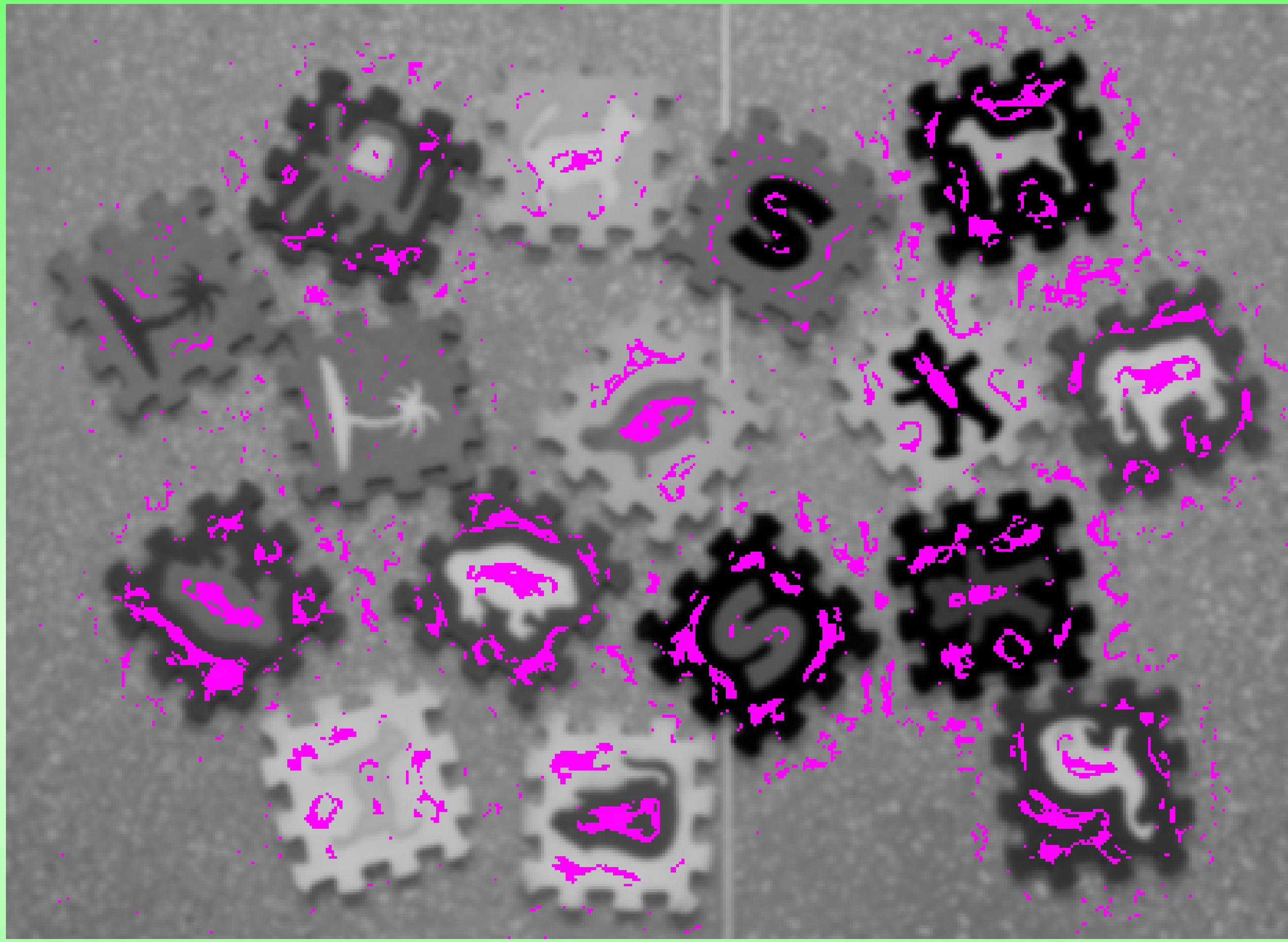
***CisCorr* is the best correlation of circular samplings C_Q and C_A:**

$$\mathrm{CisCorr}_{A,Q}(x, y) = \underset{i=0}{\overset{n-1}{\mathrm{MAX}}}\left[\left|\mathrm{Corr}(C_Q[i], C_A[x, y])\right|\right]$$

A pixel (x,y) is a first grade candidate pixel if $\mathrm{CisCorr}_{A,Q}(x, y) \ge t_1$.

The probable scale *CisPS* is the scale of the best correlation:

$$\mathrm{CisPS}_{A,Q}(x, y) = \underset{i=0}{\overset{n-1}{\mathrm{ARGMAX}}}\left[\left|\mathrm{Corr}(C_Q[i], C_A[x, y])\right|\right]$$



The output of Cifi, with first grade candidate pixels in magenta. Each candidate pixel has an associated probable scale.

4. Radial sampling filter (Rafi)

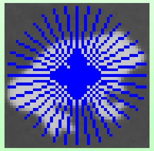
Rafi uses the projections of *A* and *Q* on radial lines to upgrade some first grade pixels to the second grade.

Radial sampling *Ras* is the average grayscale of the pixels of image *B* on a radial line:

$$\mathrm{Ras}_B^{\lambda}(x, y, \alpha) = \int_0^{\lambda} B(x + t\cos\alpha, y + t\sin\alpha)\, dt$$

Vector R_Q of radial samplings of *Q* is:

$$R_Q[j] = \mathrm{Ras}_Q^{r_l - 1}(x_0, y_0, \alpha_j), \quad 0 \le j < m$$



At each first grade pixel (x,y), *A* is radially sampled at its probable scale s_i:

$$R_A[x, y, j] = \mathrm{Ras}_R^{s_i r_l - 1}(x, y, \alpha_j), \quad 0 \le j < m \text{ and } (x, y) \in \mathrm{f\_gr\_cand}(A)$$

RasCorr **is the largest correlation between the radial samplings** $\boldsymbol{R_Q}$ **and** $\boldsymbol{R_A}$**:**

$$\mathrm{RasCorr}_{A,Q}(x,y) = \underset{j=0}{\overset{m-1}{\mathrm{MAX}}}\left[\left|\mathrm{Corr}\left(R_A[x,y], \mathrm{cshift}_j(R_Q)\right)\right|\right], (x,y) \in \mathrm{f\_gr\_cand}(A)$$

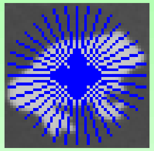
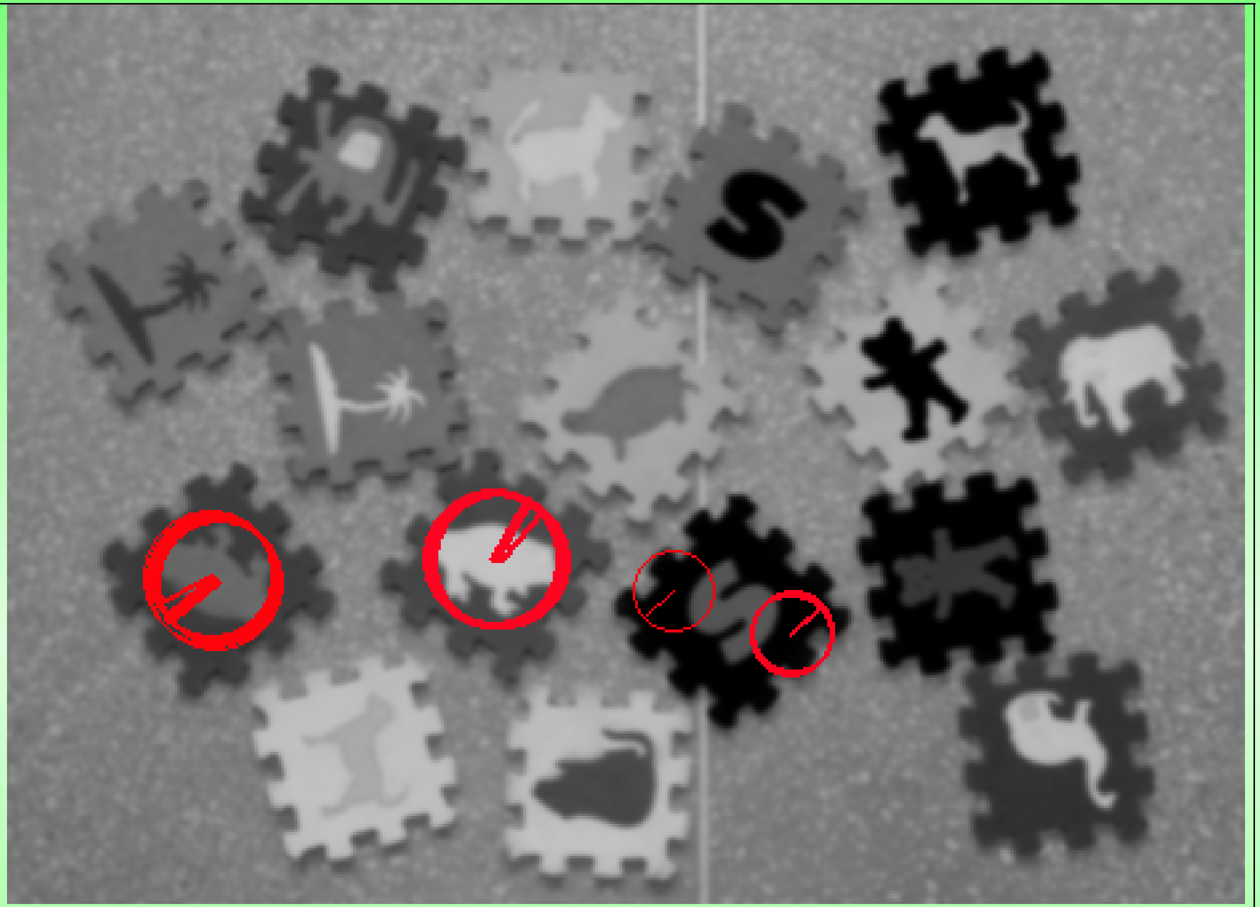
where ***cshift$_j$*** **means circular shifting** ***j*** **positions.**

A first grade pixel (x,y) is upgraded to the second grade if:

$$\mathrm{RasCorr}_{A,Q}(x,y) \geq t_2$$

The probable rotation angle at (x,y) is the angle that yields the largest correlation:

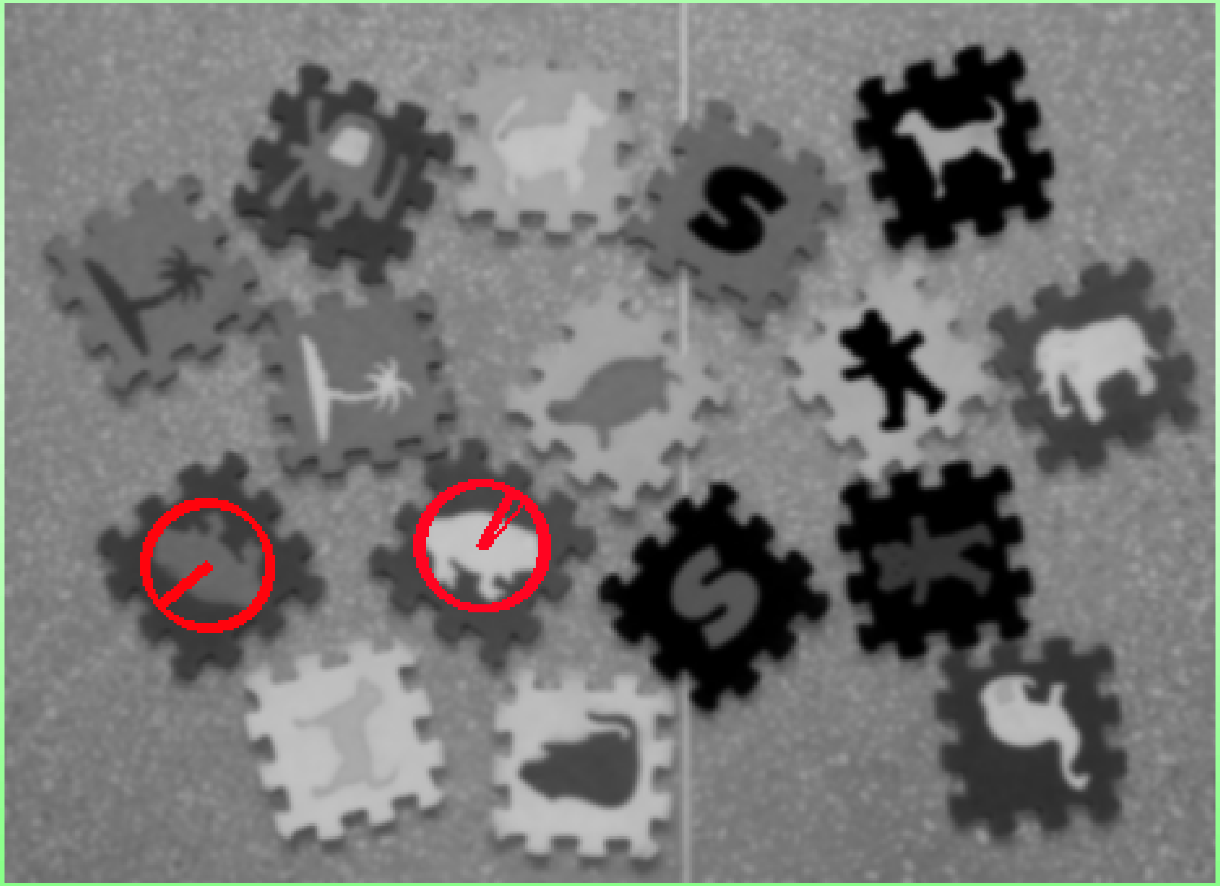
$$\mathrm{RasAng}_{A,Q}(x,y) = \underset{j=0}{\overset{m-1}{\mathrm{ARGMAX}}}\left[\left|\mathrm{Corr}\left(R_A[x,y], \mathrm{cshift}_j(R_Q)\right)\right|\right]$$

 Radial sampling lines	 The output of Rafi. Each second grade candidate pixel has an associated scale and angle.

5 Template matching filter (Tefi)

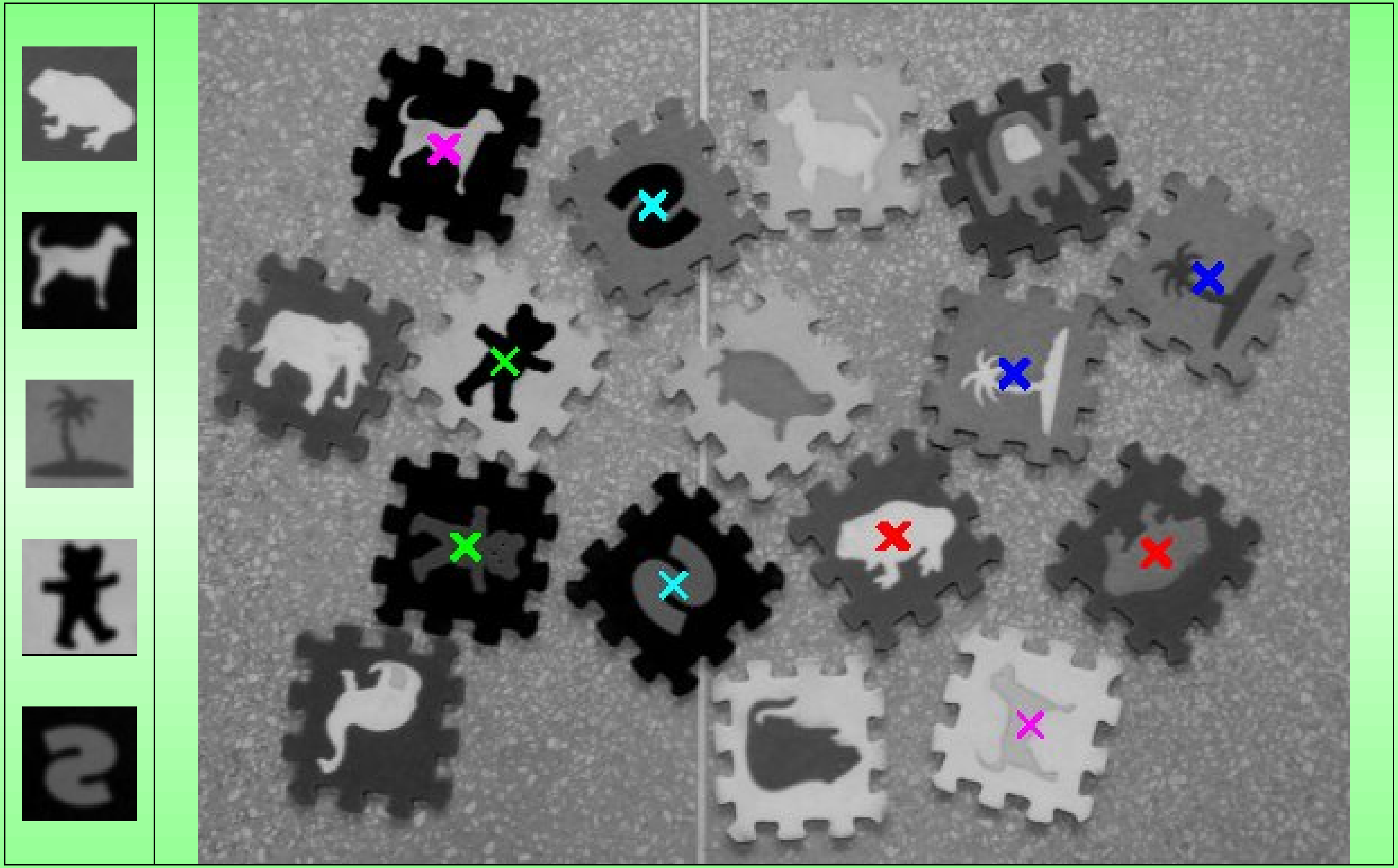
Tefi is the template matching applied at the second grade pixels, using the scales and angles determined by Cifi and Rafi.

If the correlation of the template matching is above t_3, Q is considered to be found.



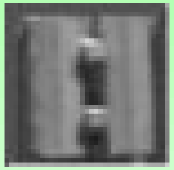
The output of Tefi.

6.1 Experiments



700 instances of 5 template toys. All matchings were perfect, without any false negative or false positive.

116 instances of McDonald’s symbols. 1 false positive and 2 false negatives.



187 instances of buildings with a specific shape. 18 false positives and 16 false negatives

6.2 Parameters

We tested the sensitivity of Ciratefi to different parameters.

The only really important parameter is t_3.

Other parameters have a large range of values that does not influence the results.

However, unappropriate parameters make the algorithm slower.

6.3 ***Preliminary*** comparison with SIFT

Sift is a well-known technique for extracting scale and rotation invariant keypoints with corresponding features.

Fair comparison is difficult because:

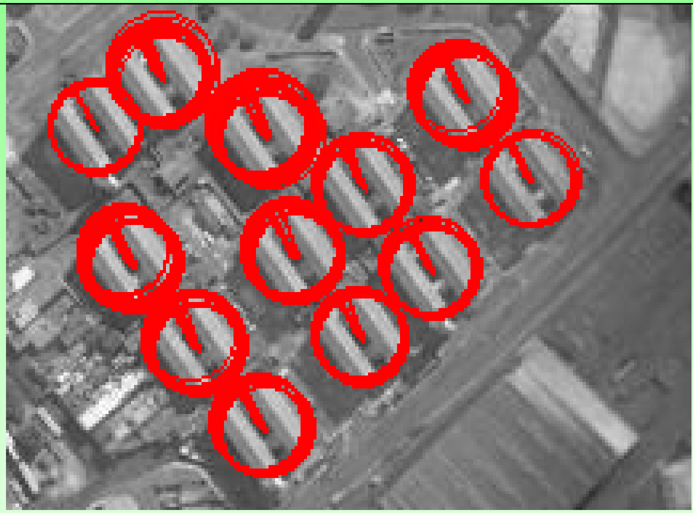
- **The primary goals of Sift and Ciratefi are different.**
- **Results depend on the parameter settings.**
- **Results depend on the chosen application.**

Preliminary **comparison results:**

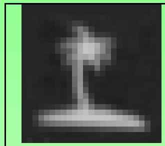
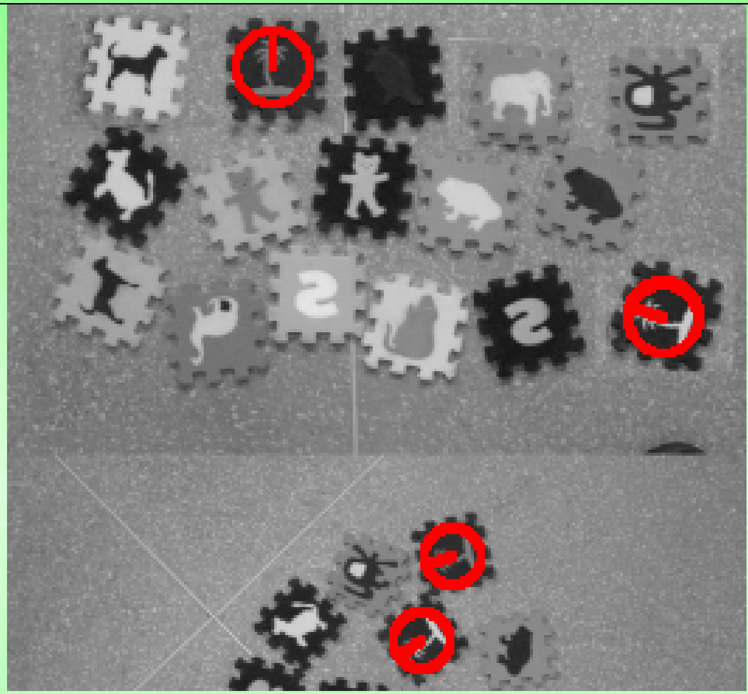
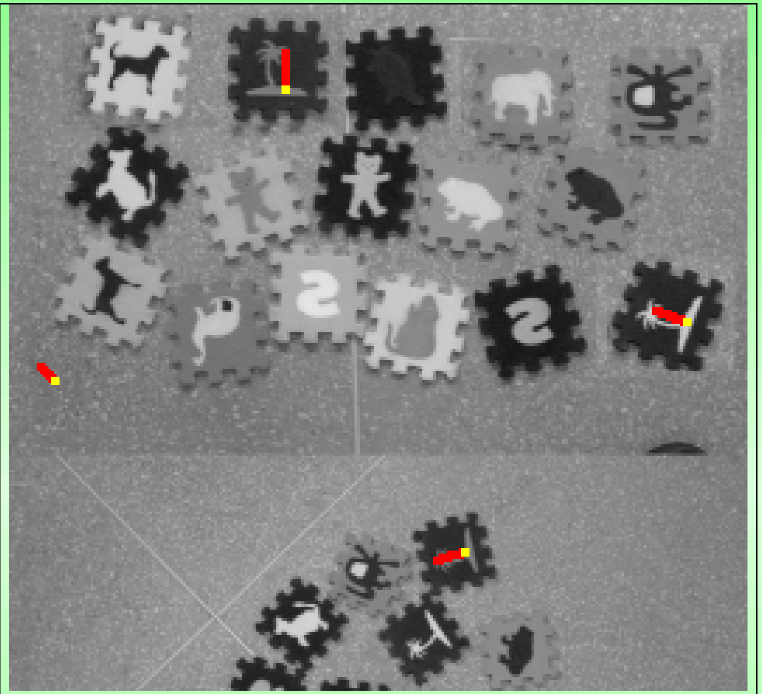
Sift is faster. To find 31×31 template in 920×1280 image, Sift takes 22s and Ciratefi takes 44s (with appropriate parameters) or more (with ill-adjusted parameters).

Ciratefi seems to be more accurate, at least for the examples we tested.

Example:

 Query	 Ciratefi: No errors.	 Sift: 1 false positive and 2 false negatives.

Example:

 **Query**	 **Ciratefi: No errors.**	 **Sift: 1 false positive and 1 false negative.**

7. Conclusions

- **We have presented Ciratefi, a new grayscale template matching invariant to rotation, scale, translation, brightness and contrast.**
- **Ciratefi is robust because it does not discard the rich grayscale information through operations like detection of edges, corners, keypoints, segmentation or binarization.**
- **Ciratefi is 800 times faster than the brute force algorithm and yields the same output.**
- **Preliminary tests show that Ciratefi is slower than Sift, but more precise, for the examples tested.**
- **Possible improvements and future works include:**
  - **Color and multispectral template matchings.**
  - **Using features different than the average grayscale.**
  - **Faster implementation.**