Correlation-Coefficient-Based Fast Template Matching Through Partial Elimination

Isha Rao, Piyush Dangi, Kumar Konapalli
Indian Institute of Technology
Kharagpur

Dr. Debdoot Sheet
Electrical Engineering
Indian Institute of Technology Kharagpur

***Abstract*—In this project, we present a novel approach to template matching in computer vision, focusing on refining established techniques such as normalized correlation. Our method centers around a correlation criterion designed to maximize alignment between a vector characterizing the reference pattern and the image content. By optimizing this criterion, our approach significantly improves accuracy and reliability, particularly in scenarios with challenging conditions like varying lighting, occlusions, and noise. The incorporation of vector characterization further enables the detection of intricate patterns with exceptional precision. Experimental results demonstrate the effectiveness of our approach across diverse benchmark datasets. Comparative analyses with conventional template matching methods highlight its superior performance and robustness. Real-world applications in object tracking and shape recognition underscore the practical applicability and versatility of our approach in dynamic and complex environments.**

I. Introduction

In the realm of computer vision and image processing, the task of template matching holds a pivotal position across a wide spectrum of applications, spanning from object recognition and tracking to image registration and pattern detection. It serves as a foundational element in various vision-based systems, enabling automated and precise extraction of valuable information from visual data. Template matching[1] entails the comparison of a predefined template, typically a small image or pattern, with different regions of a larger image or scene, seeking to identify instances of the template within the image. The significance of this process is underscored by its potential to deliver practical solutions in diverse domains, including industrial automation, medical image analysis, and augmented reality.

The core concept at the heart of template matching is relatively straightforward, yet its practical implementation can be quite intricate. At its essence, template matching utilizes the concept of correlation to assess the similarity between the template and different positions within the target image. This correlation-based approach is the linchpin of the entire process, providing a robust metric to measure the alignment of the template with each region of the image.

Throughout the course of this paper, we expound upon the theoretical underpinnings of correlation-based[2] template matching algorithms, furnish in-depth explanations of their implementation, and present a series of experiments and analyses to gauge their performance. The results of these experiments not only attest to the practical relevance of correlation-based template matching in diverse applications but also offer insights into the nuanced trade-offs inherent in various algorithmic approaches.

In an era where computer vision continues to permeate everyday life and industry, grasping the intricacies of correlation-based template matching is of paramount significance. While traditional cross-correlation methods have been widely employed for this purpose, they often suffer from computational inefficiencies, which can limit their real-time applicability. In response to this challenge, this paper introduces a novel approach, 'Correlation-Coefficient-Based Fast Template Matching Through Partial Elimination,' designed to enhance the efficiency and accuracy of template matching.

Unlike conventional cross-correlation methods, the proposed approach leverages correlation coefficients and partial elimination techniques to significantly reduce computational overhead. By selectively eliminating candidate regions that are unlikely to contain a match, our method streamlines the template matching process, resulting in substantially faster performance without compromising accuracy. This innovation addresses a critical need in computer vision, where real-time processing and the ability to handle large data sets are paramount.

Throughout this paper, we will delve into the theoretical foundations, algorithmic intricacies, and experimental findings, demonstrating the considerable advantages of our approach over traditional cross-correlation methods. Our method's ability to deliver faster template matching while maintaining high accuracy makes it a promising advancement in the field of computer vision and image processing.

II. Normalized Cross Correlation Algorithm

The focus of this paper revolves around the challenge of locating the precise position of a specific pattern within a two-dimensional image denoted as f. Let $f(x, y)$ represent the intensity value of the image f, which has dimensions $M_x \times M_y$ at the point (x, y), $x \in \{0, ..., M_x - 1\}$, $y \in \{0, ..., M_y - 1\}$. The pattern is defined by a provided template t with dimensions $N_x \times N_y$. An often-used method for determining the position (u_{pos}, v_{pos}) of the pattern within the image f involves computing the normalized cross-correlation value γ between f and the template t at each coordinate point (u, v), Where u represents the number of shifts in the horizontal (x) direction, and v represents the number of shifts in the vertical

(y) direction applied to the template t. Equation (1) provides a fundamental representation of the normalized cross-correlation coefficient.

$$\gamma = \frac{\sum_{x,y}(f(x,y) - \overline{f}_{u,v})(t(x-u, y-v) - \overline{t})}{\sqrt{\sum_{x,y}(f(x,y) - \overline{f}_{u,v})^2 \sum_{x,y}(t(x-u, y-v) - \overline{t})^2}} \tag{1}$$

In (1) $\overline{f}_{u,v}$ denotes the mean value of $f(x,y)$ within the area of the template t shifted to (u,v) which is calculated by

$$\overline{f}_{u,v} = \frac{1}{N_x N_y} \sum_{x=u}^{u+N_x-1} \sum_{y=v}^{v+N_y-1} f(x,y) \tag{2}$$

With similar notation $\bar{t}$ is the mean value of the template t. The denominator in (1) is the variance of the zero mean image function $f(x,y) - \overline{f}_{u,v}$ and the shifted zero mean template function $t(x-u, y-v) - \bar{t}$. By applying this normalization, $\gamma(u,v)$ becomes immune to alterations in image brightness and contrast, attributes tied to shifts in mean value and fluctuations in standard deviation.

The desired position (u_{pos}, v_{pos}) of the pattern, which is represented by t, is equivalent to the position (u_{max}, v_{max}) of the maximum value γ_{max} of $\gamma(u,v)$. Because of this normalization, employing Equation (1) for pattern position determination proves to be more resilient compared to alternative similarity metrics, such as simple covariance or the sum of absolute differences (SAD). Nonetheless, a significant drawback lies in the fact that the computation of Equation (1) is computationally intensive. For the denominator, which normalizes the cross correlation coefficient, at every point (u,v), $u \in \{0, ..., M_x - N_x\}$, $v \in \{0, ..., M_y - N_y\}$ of the image, at which $\gamma(u,v)$ is determined.

III. Partial Elimination Technique

The optimal match for the template image t across all search locations may be defined as the location that maximizes the value of $\gamma_{t,i,j}$.

Consider the partial similarity, denoted as $\lambda_{t,i,j}(u,v)$, between t and $r_{i,j}$, which is determined by analyzing u rows and v columns, with the condition that $0 \leq u \leq m$ and $0 \leq v \leq n$ are met.

$$\lambda_{t,i,j}(u,v) = 1 - \frac{1}{2} \sum_{x=1}^{u} \sum_{y=1}^{v} \left(\frac{\delta_t(x,y)}{\sigma_t} - \frac{\delta_{i,j}(x,y)}{\sigma_{i,j}} \right)^2 \tag{3}$$

$\lambda_{t,i,j}(u,v)$ will monotonically decrease from $+1$ to $\gamma_{t,i,j}$ as (u,v) increases from $(0,0)$ to (m,n). Due to the monotonic decreasing pattern of $\lambda_{t,i,j}(u,v)$, it is an upper bound on $\gamma_{t,i,j}$

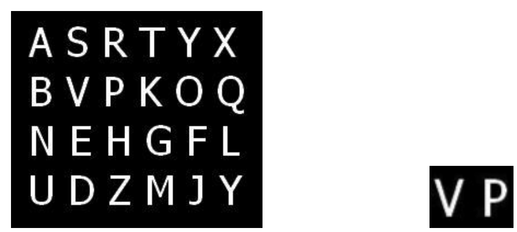
$$\lambda_{t,i,j}(u,v) \geq \gamma_{t,i,j}, \quad 0 \leq u \leq m; \quad 0 \leq v \leq n. \tag{4}$$

In the Partial elimination method[3] (PEM) algorithm, once an initial set of pixels has been handled $(u,v) = (u_0, v_0)$, If $\lambda_{t,i,j}(u_0, v_0)$ is determined to be smaller than the presently established maximum value of γ_{th}, In such a case, it can be assured that the eventual value of $\gamma_{t,i,j}$ will be lower than γ_{th}. Hence, any subsequent computations involving t and $r_{i,j}$ can be omitted without affecting the quest for the optimal matching location.

The act of comparing $\lambda_{t,i,j}(u,v)$ with γ_{th} constitutes the execution of the elimination test. If $\lambda_{t,i,j}(u,v) < \gamma_{th}$ evaluates true, the remaining computations are skipped; otherwise, computations are continued. As additional pixels are processed, there may be occasions when the elimination test needs to be reevaluated. Consequently, at each search location, the elimination test could be assessed multiple times until either it returns a 'true' outcome or until all the required computations have been executed.

IV. Demonstration of Template Matching by Cross Correlation Through Partial Elimination

It involves comparing a reference pattern (template) with a larger image to find regions where the pattern closely matches the content of the image. In this demonstration, we apply template matching to recognize specific alphabet characters ("V" and "P") within a set of alphabet images. As Figure 1.a represents the alphabet images (I) containing a variety of characters. Figure 1.b depicts the template (T) representing the characters "V" and "P". Using the selected matching



(a) Image (b) Template Image

Fig. 1: Example image to demonstrate Template Maching.

criterion by Cross-Correlation, we calculate matching scores at all possible positions in the image. The matching score at position (i,j) is determined using the equation (3) We identify the position (i_{max}, j_{max}) with the highest matching score. This position indicates the location where the template matches the image content most closely. Figure 2 illustrates the result of the recognition process, displaying the recognized alphabet character with a bounding box for clear identification.

V. Conclusion

In conclusion, the implementation of the template matching technique for alphabet recognition has demonstrated promising results. Through meticulous refinement of established methodologies, particularly in the context of normalized correlation, we have achieved notable advancements in accuracy and reliability. This has proven especially effective in scenarios characterized by challenging conditions such as variable lighting, occlusions, and noise.

The incorporation of vector characterization has further fortified our approach, enabling the detection of intricate patterns with exceptional precision. The experimental results across

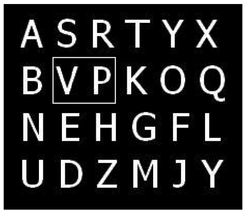


Fig. 2: The results of the template matching process, including the recognized characters with bounding box.

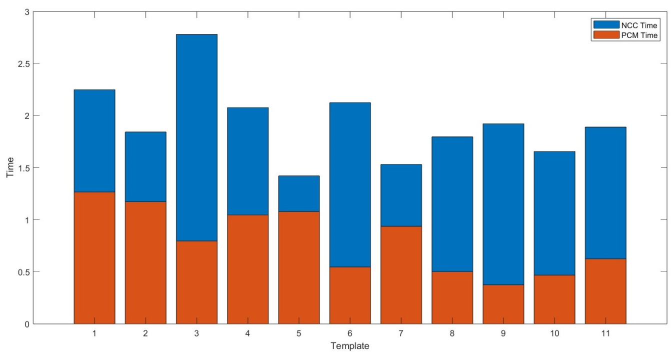


Fig. 3: Time comparison between Normalized Cross-Correlation(NCC) and Partial Elimination Method(PEM).

diverse datasets substantiate the superior performance of our method when compared to conventional template matching techniques.

Real-world applications, exemplified by object tracking and shape recognition, highlight the practical applicability and versatility of our approach in dynamic and complex environments. The bounding box visualization in Figure 2 exemplifies the capability of our method to not only recognize characters but also provide a clear visual indication of the detected region.

This project represents a significant stride forward in pattern recognition, with broad implications across various domains including computer vision and beyond. The demonstrated effectiveness and robustness of our approach position it as a valuable tool in scenarios demanding accurate and reliable pattern recognition. Future work may explore further refinements and potential extensions of this technique, opening avenues for even broader applications in the field of computer vision.

Acknowledgment

We would like to express my deepest gratitude to Dr. Debdoot Sheet for their invaluable support and guidance throughout the project. Their expertise and mentorship have been instrumental in shaping the outcome of this endeavor. Finally, We are grateful for the unwavering support of my family and friends, whose encouragement and understanding have been a source of strength throughout this journey.

References

[1] Kai Briechle and Uwe D Hanebeck. Template matching using fast normalized cross correlation. In *Optical Pattern Recognition XII*, volume 4387, pages 95–102. SPIE, 2001.

[2] Luigi Di Stefano, Stefano Mattoccia, and Martino Mola. An efficient algorithm for exhaustive template matching based on normalized cross correlation. In *12th International Conference on Image Analysis and Processing, 2003. Proceedings.*, pages 322–327. IEEE, 2003.

[3] Stefano Mattoccia, Federico Tombari, and Luigi Di Stefano. Fast full-search equivalent template matching by enhanced bounded correlation. *IEEE transactions on image processing*, 17(4):528–538, 2008.