Master Thesis Project Proposal Document.

Title - GPU cluster based

brand logo detector

in input video stream.

Company: Vion Labs AB, Stockholm

Nayan Singh Ravindra, nayan@kth.se ${\rm Saturday}~4^{\rm th}~{\rm April},~2015$

KTH Royal Institute of Technology Department of Electronic Systems Author:Nayan Singh



Page 2/6

		- 1		1 -
1 .	ΛY) T	$oldsymbol{o}$ r	${ m tts}$
\mathbf{v}	UТ	TO,	CI.	

1	Introduction	3
2	Available Methods	3



Page 3/6

1 Introduction

Pattern recognition, image search, product placement, Artificial intelligence, Recognition task in image, object detection.

What? Pattern is defined as regularity in data as perceived by humans in the form visual, auditory, haptic and olfactory cues. Human brain is well equipped in identifying patterns in their environment, it has been possible by humans to identify regularity in the data due to the presence of memory. Visual perception is one of the most interesting and powerful feature of human beings. There has been a lot of research in understanding capability of human brain in recognizing visual patterns, but still the process behind identifying complex patterns is a little understood subject in neuroscience [3].

facts?

How?

Why?

2 Available Methods

Machine learning techniques, Supervised and unsupervised learning.

• SIFT descriptor matching method: All the SIFT and SURF based methods rely upon comparing SIFT/SURF descriptors to find the best match for the queried image. The process of database creation is by using sliding window technique to compute SURF descriptors for the brand logo image. Further, during the logo detection task in an input image frame, SURF or SIFT descriptors are computed for each window of the input image. During the process the raw descriptors obtained are matched with the database to determine the identical match. The window containing the matching descriptors is expected to contain the matched logo.

Problems: Compute SURF or SIFT descriptors of the input image by using sliding window method (an exhaustive search). Although SIFT or SURF descriptors are scale, rotation and position invariant but there are highly perspective tilt dependent. For example the descriptors of affine transformed images are less likely to be matching with the non-affine transformed logo image. Since the logos can occur in various tilt perspectives in the image frame, the SIFT or SURF descriptor based matching approach is a suboptimal solution.

• Delaney triangle based method:

• Bundle minimum hashing of Bag of Visual words methodBundle Min-Hashing for Logo Recognition Romberg et.al: Method description: Stores the warped versions of the logo image in the database. Image partitioning method is used instead of sliding window method during logo detection process.

Problem: The input image frame has to be analyzed to predict the location of the logo. Localisation is a problem of its own, hence a more emphasis is given to predict the position of the objects from natural scene images. The accuracy of detection depends on the robustness of prediction of the bounding boxes around objects containing logos.

• Convolution neural network:

Logo detection techniques can be broadly categorized into various methods as follows:

• Geometry based methods: In this technique the geometric features of the logos are considered for the detection and identification of logos in the input images. Few researchers have considered this approach to accomplish the task of logo detection in images and document images. Few of the related work are explained in this section to attain better understanding of the methods and the shortcomings of the discussed geometric methods.

One of the approach to detect logos in the document images emphasize on using geometrical invariant features as in [1]. In this paper, authors focus on creating an invariant signature that uses local affine features of the logo to enable the detection process. They use euclidean invariants to capture the affine features which are further used for logo detection, feature extraction and for matching purposes [1]. The signatures obtained from this method are geometrically invariant and highly discriminative for a particular logo, which is based on shape of the logo. The proposed method is highly robust to several geometric transformations such as rotation, translation and scaling. The comparison of the signatures obtained from this method makes it more computational intensive and also the detection of multiple instances of various logos in a image frame requires sliding window technique [1].

Later in year 1998, a new method for proposed to detect logos in the scanned images based on geometrical features as in [4]. Soffer et.al in [4] used the concept of identifying logos by using the negative shape features for similarity matching. The main idea of this approach is to enable the detection of shapes by computing the positive and negative areas of the logo in the scanned image. Further



Page 5/6

the logo is represented as feature vectors containing both global and local shape descriptors. The global geometrical feature descriptors are first invariant moment, circularity, eccentricity and other rectangular properties of the logo shape are computed [4]. Also local shape descriptors such as horizontal gaps per total area, vertical gaps per total area and the ratio of negative area to the total area are computed. The similarity matching technique involves of finding the near similar logo among various classes of logos in the database. The shortcoming of this approach is that the feature vector computing algorithm requires for the input images to be bounded by rectangles prior to detection process, hence the task of logo localization and identification becomes an exhaustive task.

One more approach based on geometrical features for the detection of logos in video stills was suggested by Hollander et.al [2] in year 2003. Authors chose to address the problem of detecting logos in video stills using template matching of strings obtained from the input video stills and original logo pictures. The assumptions made by the authors were that the logos will always be embedded in the homogeneously colored background such as sports event videos and the logos are bright colored and easily recognizable from the background [2]. The process of logo identification involves finding of homogeneously colored region of pixels on the HSV representation of video stills, to detect black and white regions. The approach of detection uses sliding window technique of a heuristically chosen sizes for hue values to enable detection of all colors and intensity values [2]. Involving intensity attribute of the image still makes the detection process to be highly susceptible to lighting conditions. Further, a morphological operation of closing in applied to the obtained image and white regions that have an area greater than the threshold are chosen to perform sobel edge transform to obtain the boundaries of the edges [2]. The logo classification task involves identification of the line that covers the maximum pixels across the logo image using hough transform, the obtained line is invariant to position but variant to scale transformations. This string is unique for the logo to be classified, hence the template matching method of strings is used to find the identical logo from the database. This approach has serious shortcomings due to varied light conditions, shapes of the logo, partial occlusion and other affine transforms of the logo in the natural scenes.



References

- [1] David S Doermann, Ehud Rivlin, and Isaac Weiss. "Logo recognition using geometric invariants". In: Document Analysis and Recognition, 1993., Proceedings of the Second International Conference on. IEEE. 1993, pp. 894–897.
- [2] Richard JM den Hollander and Alan Hanjalic. "Logo recognition in video stills by string matching". In: *Image Processing*, 2003. *ICIP 2003. Proceedings. 2003 International Conference on*. Vol. 3. IEEE. 2003, pp. III–517.
- [3] Pawan Sinha. "Recognizing complex patterns". In: nature neuroscience 5 (2002), pp. 1093–1097.
- [4] Aya Soffer and Hanan Samet. "Using negative shape features for logo similarity matching". In: *Pattern Recognition*, 1998. Proceedings. Fourteenth International Conference on. Vol. 1. IEEE. 1998, pp. 571–573.