Pattern Recognition Letters 78 (2016) 36–40

Contents lists available at ScienceDirect

Pattern Recognition Letters

journal homepage: www.elsevier.com/locate/patrec

Symmetric stability of low level feature detectors☆

Craig Henderson*, Ebroul Izquierdo

Multimedia and Vision Lab, School of Electronic Engineering and Computer Science, Queen Mary University of London, Mile End Road, London, E1 4NS, United Kingdom

ARTICLE INFO

Article history:
Received 25 August 2015
Available online 9 April 2016

Keywords:
Feature matching
Feature detectors
Bilateral symmetry

ABSTRACT

We investigate the capability of low level feature detectors to consistently define feature keypoints in an image and its horizontally reflected (mirrored) image. It is our assertion that this consistency is a useful attribute of a feature detector and should be considered in assessing the robustness of a feature detector. We test ten of the most popular detectors using a popular dataset of 8677 images. We define a set of error measurements to help us to understand the invariance in keypoint position, size and angle of orientation, and we use SIFT descriptors extracted from the keypoints to measure the consistency of extracted feature descriptors. We conclude that the FAST and CenSurE detectors are perfectly invariant to bilateral symmetry, *Good Features to Track* and the Harris Corner detector produce consistent keypoints that can be matched using feature descriptors, and others vary in their invariance. SIFT is the least invariant of all the detectors that we test.

1. Introduction

There are many feature detectors documented in the literature, and used in research and practical applications to fulfil the common need to identify interest points within an image. Information at these positions can then be extracted into a descriptor and used for correspondence matching in image retrieval and classification, image alignment, image stitching, and many other applications. The two stages are often combined into one in discussion, but each are independent and the algorithms used in each can often be interchanged.

Most popular and useful *feature descriptors* are invariant to scale and rotation, and matching features from two images where they appear at different sizes or are rotated can still be successful. Invariance to bilateral symmetry in *feature detectors*, however, is less well documented. We describe our interest in this invariance and investigate the property for some popular feature detectors, assessing their consistency in finding interest points within an image and a horizontal reflection of the image. Our goal is to identify which popular feature detectors are most invariant to bilateral symmetry, and what degree of error exists in the interest point position, size and orientation.

To the best of our knowledge, no assessment of low level feature detectors with respect to their invariance to bilateral symmetry has previously appeared in the literature. The main contributions in this paper are:

- We introduce five measurements of error that we show to be useful in determining the invariance to bilateral symmetry of a feature detector; *mean distance error*, *mean size error*, *mean angle error*, *mean descriptor distance error* and the *mean descriptor match error* (Section 4).
- We measure the accuracy of bilateral keypoint position, size and angle of orientation in an established dataset [5] of 8677 images (Section 5).
- We evaluate the capability of popular detectors to find consistent interest points (Section 6).

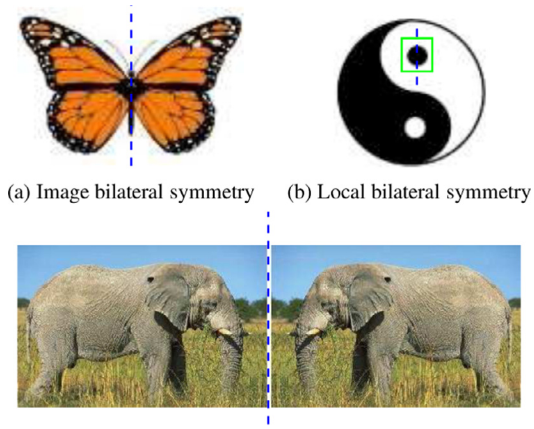
2. Bilateral symmetry

Bilateral symmetry describes a symmetry through a vertical plane in an image, and can occur at different scales. Fig. 1 shows two examples; (a) the image as a whole is bilaterally symmetrical because the right hand side of the plane (the dotted blue line down the centre of the image) is a mirror image of the left hand side and (b) the highlighted section of the image is bilaterally symmetrical although the image as a whole is not. Detected keypoints in an image are generally very small and detection of bilateral symmetry will be at a finer scale than both of these examples. Fig. 1c shows our test case where we horizontally mirror the image to assess inter-image bilateral symmetry, and Fig. 2 shows a real life example from a London street CCTV camera that demonstrates the need of reflection invariance in analysing CCTV images.

☆ This paper has been recommended for acceptance by Dr. J. Laaksonen.

* Corresponding author. Tel.: +44 7584430212.
E-mail address: c.d.m.henderson@qmul.ac.uk (C. Henderson).

http://dx.doi.org/10.1016/j.patrec.2016.03.027
0167-8655/



(a) Image bilateral symmetry (b) Local bilateral symmetry

(c) Inter-image bilateral symmetry

Fig. 1. Bilateral symmetry (mirror reflection) at different scales.

Fig. 2. Motivating example. Cropped frames from a CCTV camera capture images of a man wearing a Nike hoodie (left) and later in the video having turned the hoodie inside-out, showing the Nike logo in reverse.

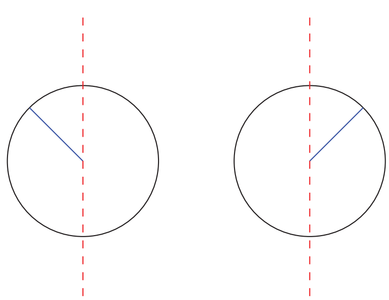


Fig. 3. Reflecting a keypoint. (For interpretation of the references to colour in this figure, the reader is referred to the web version of this article.)

A man wearing a hoodie exhibiting a Nike sportswear logo is later captured wearing the hoodie inside-out, with the Nike logo in reverse. Consistency in the detected position of a keypoint between an image and its horizontal reflection is important to enable a reflection invariant descriptor such as MIFT [9] to be extracted from the same point in the logo to maximise the potential to achieve a correspondence.

A keypoint is defined by its (x, y) co-ordinates, size, and sometimes, angle of orientation (Fig. 3). We reflect a keypoint in its centre line (red dotted line). Let I be the image in which the keypoint is found, and I_x be the x-dimension of the image; the width. Let α be the angle of orientation, measured clockwise from 0° parallel to the x-axis. Then, the new values for the x position of the keypoint is x' and the new angle of orientation is α', thus

$$x' = I_x - x - 1 \tag{1}$$

$$\alpha' = \pi - \tan^{-1}\left(\frac{-\sin\alpha}{-\cos\alpha}\right) \tag{2}$$

3. Feature detectors

Feature detectors are used extensively in all areas of computer vision to identify parts of an image which contain pixel information that can be useful in many applications. Numerous detector methods have been described in the literature, and many have become popular for different tasks. Two distinct categories of feature detectors exist; *keypoint* detectors and *region detectors*. Recent trends in Deep Learning use features that are discovered automatically during the training process. In this paper, we concentrate on low-level feature detectors that can be described algorithmically, and how these perform in respect to reflection invariance.

Rosten and Drummond [19] learn a ternary decision tree that can detect points with high repeatability, to create FAST; *Features from Accelerated Segment Test*. The BRISK detector [12] extends FAST with *an assembly of a bit-string* descriptor from intensity comparisons retrieved by dedicated sampling of each keypoint neighbourhood. ORB [20], is also based on the FAST detector from where the name is derived *Oriented FAST and Rotated BRIEF*, where BRIEF [4] is a feature descriptor. The Harris Corner detector (HARRIS, [10]) is a combined corner and edge detector based on the local auto-correlation function, and was extended by Shi and Tomasi [21] to create *Good Features to Track* (GFTT). *Scale Invariant Feature Transform* (SIFT, [13]) is perhaps one of the most well known and commonly used detectors and uses a histogram of local oriented gradients that are measured in a pyramid of Gaussians to achieve scale invariance. *Speeded-Up Robust Features* (SURF, [2]) is a faster SIFT-inspired detector, using Hessian matrix to achieve good performance in computation time and accuracy. The final keypoint detector we evaluate is CenSurE [1], described as a fast variant of the upright SURF descriptor, and sometime called STAR.

In addition to keypoint detectors, we use two related region detectors from the same primary author; *Maximally Stable Extremal Regions* (MSER, [7]) for grey-scale images and *Maximally Stable Colour Regions* (MSCR, [6]).

4. Experiments and data

We assess the eight keypoint detectors and two region detectors described above, using the well established CALTECH101 dataset [5]. The dataset consists of 8677 JPEG images grouped into 101 categories, and contains a variety of image styles including cartoons and photographs of objects, human faces, animals and natural scenes. MSCR is the only detector that works with 3-channel colour images and for all other detectors, the original colour images are first converted to intensity images.

To measure the reflection invariance of the detectors with respect to bilateral symmetry, we use SIFT descriptors and measure their distance in feature space. Feature descriptors are themselves not invariance to bilateral symmetry and descriptors from an original image cannot be compared to a corresponding feature in a mirrored image. To overcome this, we extract feature descriptors from the original image using the detected keypoint attributes, and from reflected attributes detected in the mirror image. Let I represent an original image and M be the mirror of I. Then K_I and K_M represents keypoints detected in each of I and M respectively. Keypoints K_M are reflected as $k'_M \in K'_M$ using Eq. (1) and (2), and feature descriptors are extracted from I using K'_M.

Our assessment is based on keypoint size and position. For features found by region detectors, we define a keypoint at the centre of the non-orthogonal (rotated) bounding rectangle of the region, and measure the size of the region as the encasing circle.