

Review On Different Feature Extraction Algorithms

Shilpa G. Sonone¹, Dr. Mahadev Kokate²

Department of Electronics and Telecommunication Engineering

K.K.Wagh Institute of Engineering Education & Research

Nashik 422003, INDIA

email:

¹shilpasonone04@gmail.com,

²mdkokote@kkwagh.edu.in

Abstract -> In visual sensor networks, energy conservation is main issue because more energy is consumed in processing and transmitting the large amount of data. The traditional approach, denoted here in after as "Compress-Then-Analyze" (CTA), consists in the following steps: the visual content is acquired by a sensor node in the form of still images or video sequences; then, it is encoded and efficiently transmitted to a central unit where visual feature extraction and analysis takes place. In this process large amount of energy is consumed in compression processes and transmission processes also transmitting data at low bandwidth is also a main issue so to overcome this issues new approach is used i.e. "Analyze-Then-Compress" (ATC) In particular; nodes process visual content in order to extract relevant information in the form of visual features. Then, such information is compressed and sent to a central unit, where visual analysis takes place. Here In this paper we just overview the different compression techniques or different algorithms in real valued (SIFT And SURF) and binary feature extraction (BRIEF, BRISK, Bin Boost) and also describe coding technique in WWSN.

Key Words: Visual features, binary descriptors, SIFT, SURF, BRIEF, BRISK, Bin-Boost, video coding.

1. INTRODUCTION

Visual features provide a succinct, yet efficient, representation of the underlying visual content, which is robust and invariant to many global and local transformations. They are effectively employed in many tasks, ranging from image/video retrieval, object recognition, object tracking, image registration, structure-from

motion, etc.[1] Visual feature extraction algorithms consist of two main components: the detector, which identifies salient keypoints within an image; and the descriptor, which provides a concise representation of the image patch surrounding each keypoint. There are different descriptor are designed till now which perform same processes such as namely pre-smoothing, transformation and spatial pooling.[2]

In WWSN different compression techniques are used to compressed data and transfer it to the base station. The tradition approach is "Compress-Then-Analyze"(CTA)is based on the following steps: the signal of interest (i.e., a still image or a video sequence) is acquired by a sensor node, then it is compressed(e.g., resorting to JPEG or H.264/AVC coding standards) in order to be efficiently transmitted over a network. Finally, visual analysis is performed at a sink node [3]- [6].This type is used in different application like video surveillance, smart cameras, etc. but it represent a lossy visual features and having bandwidth limitation so we switch toward new technique Analyze-Then-Compress.

An alternative technique to the Compressed-then-Analyze is Analyze -then-Compressed is gaining popularity in the research community. In this process local features are extracted directly from sensing node. Then, they are compressed to be efficiently dispatched over the network. There are different algorithms are used for extraction of local features such as SIFT, SURF, BIREF, BRISK, Bin-Boost etc.

2. RELATED WORK

Compression of visual features is one of the important task in WWSN. The main problem related to transmission of date is bandwidth, in WWSN the date which is transferred to the sink node having low bandwidth. Several works tackled this problem for the case of features extracted from still images, proposing methods to efficiently encode state-of-the-art visual features [7]or to modify the design of local feature

extraction algorithms, so that the representation of the underlying visual content is more suitable for compression [8]. In this context, an ad-hoc MPEG group on Compact Descriptors for Visual Search (CDVS) is currently working towards the definition of a standard tailored to this scenario [9].

Processing of the visual features require more data so the energy require for transmission of the large is also more so this is also one of the problem regarding transmission of data directly to the base station[10]. So different algorithms are used to extract visual features and then designed encoder which encodes the data at low computational complexity.

First, Real-valued features such as SIFT [11] or SURF [12], These are methods for extracting distinctive invariant features from images that can be used to perform reliable matching between different views of an object or scene. The features are invariant to image scale and rotation distortion. This also describes an approach to using these features for object recognition. The recognition proceeds by matching individual features to a database of features from known objects using a fast nearest-neighbour algorithm, followed by a Hough transform to identify clusters belonging to a single object, and finally performing verification through least squares solution for consistent pose parameters. This approach to recognition can robustly identify objects among clutter and occlusion while achieving near real-time performance.

BRIEF [13] propose to use binary strings as an efficient feature pointed scriptor, which we call BRIEF. We show that it is highly discriminative even when using relatively few bits and can be computed using simple intensity difference tests. Furthermore, the descriptor similarity can be evaluated using the Hamming distance, which is very efficient to compute. Thus this algorithm is suitable whenever energy resources area issue, such as in the case of low-power devices, where they constitute the only available option. The processing pipeline for the extraction of local features comprises: i) a keypoint detector, which is responsible for the identification of a set of salient keypoints within an image, and ii) a keypoint descriptor, which assigns a description vector to each identified keypoint, based on the local image content. Within the class of local binary descriptors, BRIEF [14] computes the descriptor elements as the result of pairwise comparisons between (smoothed) pixel intensity values that are randomly sampled from the neighbourhood of a keypoint.

BRISK [14], FREAK [15] and ORB [16] are inspired by BRIEF, and similarly to their predecessor, are also based on pair wise pixel intensity comparisons. They differ from each other in the way pixel pairs are spatially sampled in the image patch surrounding a given keypoint.

BAMBOO [17] adopts a richer dictionary of pixel intensity comparisons, and selects the most discriminative ones by means of a boosting algorithm. This leads to a matching accuracy similar to SIFT, while being 50x faster to compute. Bin Boost [18], which proposes a boosted binary descriptor based on a set of local gradients. Bin Boost is shown to deliver state-of-the-art matching accuracy, at the cost of a computational complexity.

In this paper we just overview the different algorithms of extracting features or compare different types of descriptors for video coding. Now in next part we explain proposed system of video coder by using BRISK algorithm, then proposed flow chart of video coder and BRISK algorithm, and finally conclusion.

3. PRAPOSED SYSTEM OF VIDEO CODER

As described in previous section different compression techniques are used in WWSN and different algorithms are used for visual feature extraction. Here we describe video coder in which features extracted from video by using binary feature extraction algorithm i.e. BRISK algorithm. After extraction of features by BRISK algorithm then it coded by using entropy coding as shown in figure 1. then this data is transferred wireless channel to the base station. Here descriptors are coded to meet the energy and bandwidth requirement

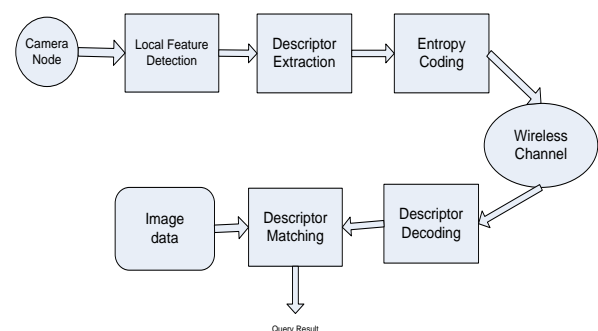


Fig.1 Proposed Block Diagram for video coding and feature matching.

