Distributed Video Surveillance Using Smart Cameras

Most of the current commercial video surveillance systems rely on a classical client/server architecture to perform face and object recognition. In order to support the more complex and advanced video surveillance systems proposed in the last years, companies are required to invest resources in order to maintain the servers dedicated to the recognition tasks.

Automated area surveillance addresses the real time observation of people and objects in a busy environment, leading to their recognition and description of their actions and interactions.

Human resources used in  
the field of the video surveillance services are both  
costly and not reliable.

The most important issues include  
tasks like people detection, recognition and tracking;  
object recognition, motion analysis, etc

(Motivation)Executing the needed image  
processing tasks, for offering this kind of automated  
video surveillance in real-time, requires significant  
storage and processing resources. The need to process big amount of video data represents a common  
bottleneck of the automated video surveillance systems, especially when the number of camera providing  
video streams and the number of faces to be simultaneously recognized and tracked is large

One of the straightforward approaches for resource  
allocation in video surveillance services is to use  
client-server model of communication. In other words,  
the surveillance devices stream the video directly to a  
main powerful server, where the data can be analyzed  
[23]. Despite of the effectiveness and simplicity of  
such architecture, this approach presents several negative sides. First, relying all the functionality on a single  
server creates a bottleneck for the system security and  
reliability. A high rate of the recognition tasks on theserver can lead to delays in the image processing. Second, in order to analyze all the video streams coming  
from the surveillance devices the target organization  
needs to maintain big and costly infrastructure of  
servers dedicated only to the surveillance task. Moreover, such a service become too costly for small and  
medium size business, which are not able to allocated  
so much financial resources only for the surveillance.

This work extends and enhances our previously  
published research work [16], where we have proposed a distributed algorithm for load balancing  
between Smart Sensing Units for video surveillance  
task. The adaptive algorithm distributes, at run time,  
the recognition tasks between the resources of surveillance devices and servers. The detection and recognition tasks are executed locally by surveillance devices,  
which exploit both the spatial and temporal topology  
of the moving people, to cache and reuse locally parts  
of the classification features. Only when devices are  
not able to execute the recognition task with the cache,  
a recognition request is sent to the server. We extend  
the previously proposed adaptive algorithm considering the overhead of real classification techniques. In  
our evaluation we implemented and tested classification  
algorithms for face recognition: 1NN and the weighted

In this paper, we consider the following classifiers  
for face recognition: the first nearest neighbor (1NN)  
and the weighted k-nearest-neighbor (weighted kNN)  
implemented with LBP

In this paper, we propose a distributed camera-aided  
protocol for area video surveillance based on image  
classification. In order to minimize the classification  
load on the main server, recognition tasks take place  
on the SSUs when possible. To perform person recognition, a SSU uses the local resources together withthe resources of the neighbor SSUs. The surveillance  
devices fall back to the main server when the classification cannot be done in the desired time interval.  
We have evaluated two different classification algorithms (1NN and weighted kNN) implemented with  
LBP. Among the considered classifiers the best performance results are obtained by the weighted kNN  
algorithm. Therefore, we have evaluated the load of  
MS and network characteristics based on weighted  
kNN classifiers applied on SSUs. Our evaluation simulated the surveillance area by considering a number  
of hotspots. The bandwidth consumption of the SSUs  
that are responsible for the surveillance of the hotspots  
resulted much higher than for the other SSUs. This  
suggests to use more powerful devices for the hotpots  
surveillance, or to apply a combined approach in which  
the system switches to a pure client-server model recognition and streams video directly to the main server  
for processing in case of high people concentration at  
one surveillance spot. Moreover, a further optimization of the bandwidth consumption could be achieved  
by applying algorithms for the people trajectory estimations directly inside the SSUs, in order to reduce  
the features broadcasting between neighbor SSUs.  
Finally, the evaluation shows that partially placing  
the recognition tasks on the local resources of surveillance devices can reduce the load on the main server  
up to 50%. As a consequence, it is possible to run a  
surveillance service by employing less powerful centralized servers, or, alternatively, a large server can be  
re-used for the other tasks in the less busy hours.

Face detection allows a computer to identify human faces in digital images and video. This task is not trivial due to the fact that human face in video frames is usually a dynamic object, with high degree of variability in its appearance. Different illumination conditions, facial poses and expressions, occlusion, rotation can affect a face detection algorithm.