Automatic Classification of Video Sequences into Specified Generalized Use Classes of Target Size and Lighting Level

Mikołaj I. Leszczuk Marcin Witkowski

Department of Telecommunications
AGH University of Science and Technology
Kraków, PL-30059

June 11, 2012

Table of contents

1. Introduction
2. Framework for Describing Public Safety Video Applications
3. Classification of Video Sequences into Specified Generalized Use Classes
   - Research
   - Analysis of Results on Target Size
   - Analysis of Results on Lighting Level
4. Methods for Automatic Classification of Entire GUC Sequences
   - Data preparation
   - Target Size Classifier
   - Lighting Level Classifier
5. Method Evaluation
6. Conclusions and Further Development
7. Acknowledgements

Introduction

- Transmission and analysis of video frequently used for variety of applications outside entertainment sector, to perform specific tasks
  - Security
  - Public safety
  - Remote command and control
  - Tele-medicine
  - Sign language
- Each application consisting of some type of recognition task
- Different QoE for entertainment and recognition tasks videos
- Video Quality in Public Safety (VQiPS) Working Group, est. 2009 by DHS, developing user guide for public safety video applications
- The approach taken by VQiPS:
  - Not attempting to address each of public safety video applications
  - Remaining application-agnostic and basing on common features

Five Parameters Impacting Ability to Achieve Recognition Task, Selected as Being of Particular Importance

- **Usage time-frame** — specifying whether video to be
  - Analysed in real-time
  - Recorded for later analysis
- **Discrimination level** — specifying fine level of detail sought from video
- **Target size** — specifying anticipated Region Of Interest (ROI) in video to occupy relatively small or large percentage of frame
- **Lighting level** — specifying anticipated lighting level of scene
- **Level of motion** — specifying anticipated level of motion in scene

Representation of Determination Process of Generalised Use Classes (GUCs) Formed from Referred Parameters

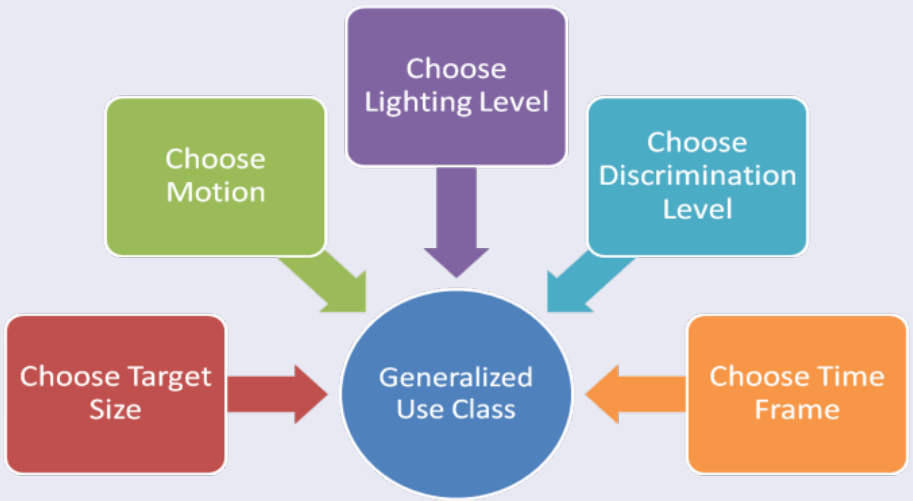
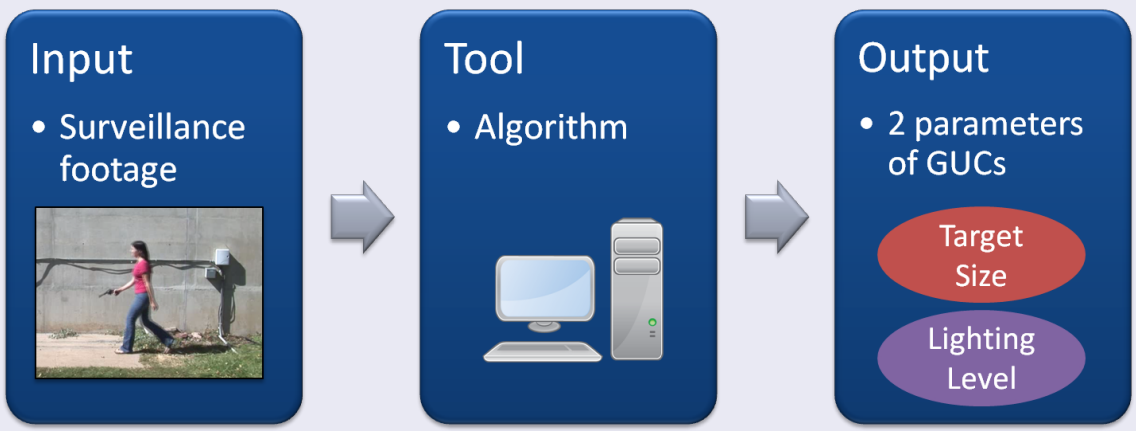


Figure: Classification of video into generalized use classes as proposed by VQiPS

Classification of Video Sequences into Specified GUC

- **Objective** – to develop tool that would automatically classify input sequence into one of GUCs
- **Challenge** – GUC description not defining particular characteristics of targets, usable as criterion for automatic algorithms
- **Parameters** – as NTIA conducted research on motion level, AGH approached remaining parameters: target size and lighting level



Work Description – Block Diagram

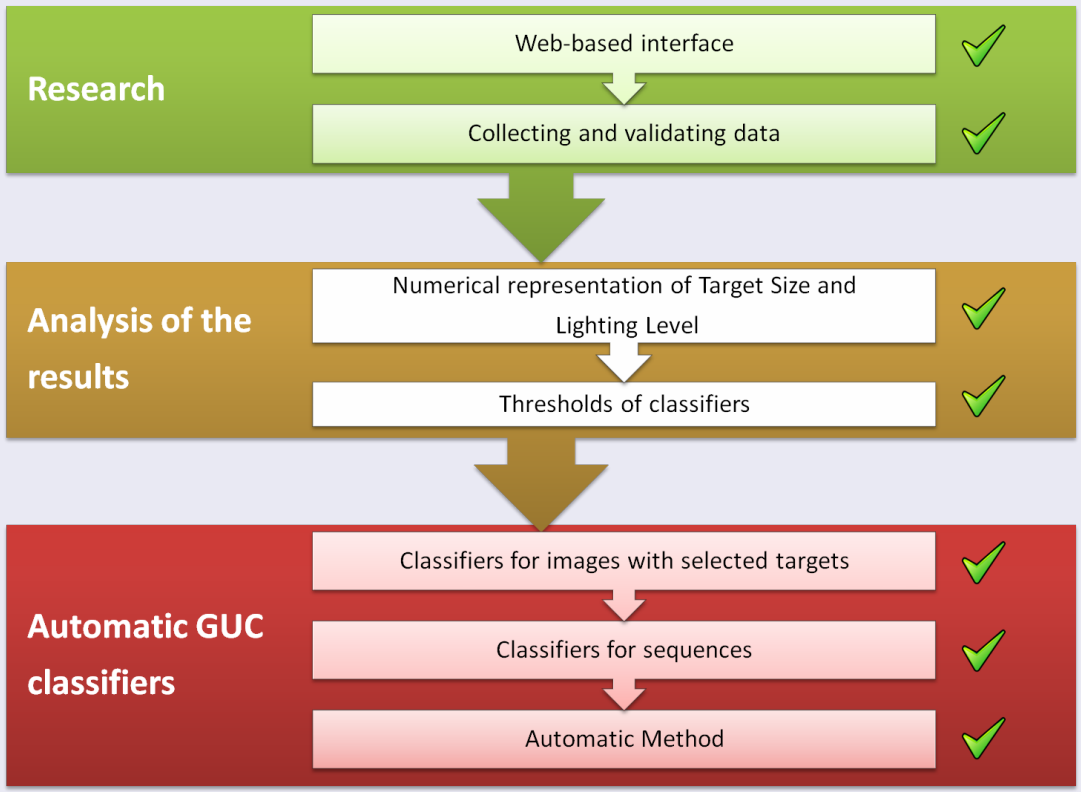


Figure: Working on automatic classification into GUCs

Table of contents

1. Introduction
2. Framework for Describing Public Safety Video Applications
3. Classification of Video Sequences into Specified Generalized Use Classes
   - Research
   - Analysis of Results on Target Size
   - Analysis of Results on Lighting Level
4. Methods for Automatic Classification of Entire GUC Sequences
   - Data preparation
   - Target Size Classifier
   - Lighting Level Classifier
5. Method Evaluation
6. Conclusions and Further Development
7. Acknowledgements

Assumptions of Research Tool

Functionalities:

- Watching video samples
- Selecting targets by drawing on frames and describing them
- Selecting lighting level of whole sequence and particular targets

Features:

- Intuitive
- Easily accessible
- Well performance at most popular web browsers

Developed Tool

Figure: Outlook of interface

Results

The set of answers consisted of 616 target selections. Preparation for analysis:

- Manual validation as a result of subjective character of the test
- Excluded entries contained:
  - actions
  - two or more targets selected at once
  - no particular target selected
  - the same target selected more than once by one end-user
- Finally we have got **553** valid answers.

Results — Examples of Excluded and Validated Entries

Figure: From top: action, many targets at 1 selection, no particular target

Results — Grouping targets

- Commonalities between selections and descriptions
- Conditions
  - Common 66.7% ($\frac{2}{3}$) of size selections and descriptions
  - Target was selected at least twice

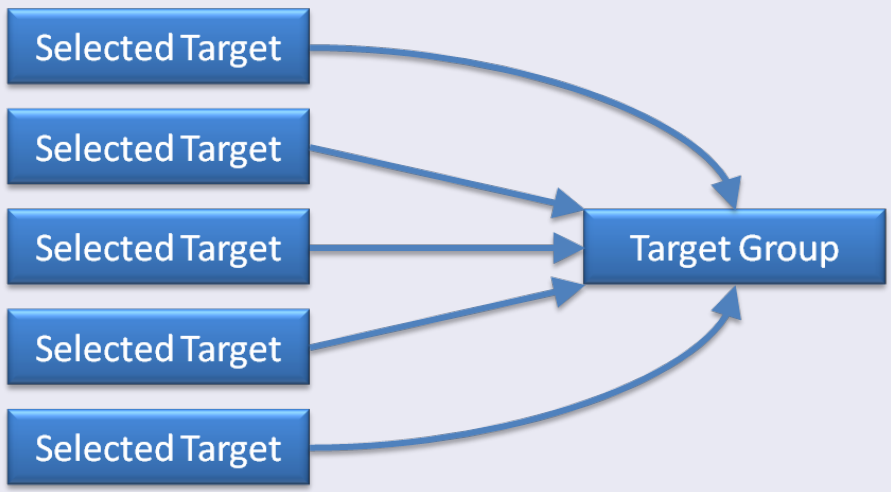


Table of contents

1 Introduction
2 Framework for Describing Public Safety Video Applications
3 Classification of Video Sequences into Specified Generalized Use Classes
- Research
- Analysis of Results on Target Size
- Analysis of Results on Lighting Level

4 Methods for Automatic Classification of Entire GUC Sequences
- Data preparation
- Target Size Classifier
- Lighting Level Classifier

5 Method Evaluation
6 Conclusions and Further Development
7 Acknowledgements

Target Size

- VQiPS defining 2 sizes of anticipated ROIs (targets)
  - Small
  - Large
- Finding binary classification criterion based on subjects
- Different numerical metrics of target sizes calculated
  - $F1$ — F1 score
  - A — Measuring accuracy
  - P — Precision
  - R — Recall
- $TS = \frac{max(x,y)}{X \veebar Y}$
  - TS — Target Size metric
  - x, y — Size of selected ROI
  - $X \veebar Y$ — Respective length of frame dimension
- $A_{max}(TS = 40\%) \geq 85\%$

Target Size — Histogram

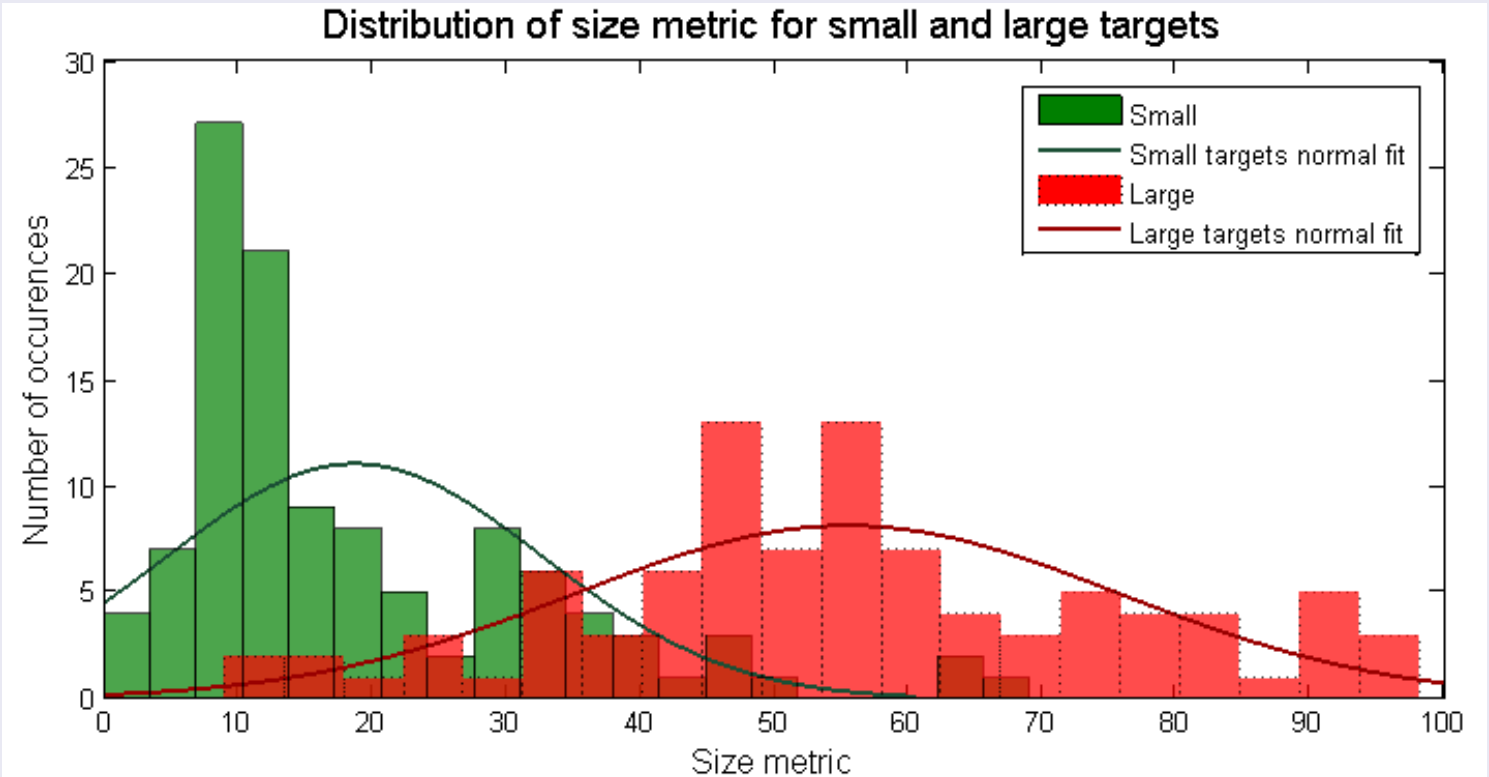


Figure: End-user selections of target size in function of size metrics

Target Size — Various Metric Values for Statistics

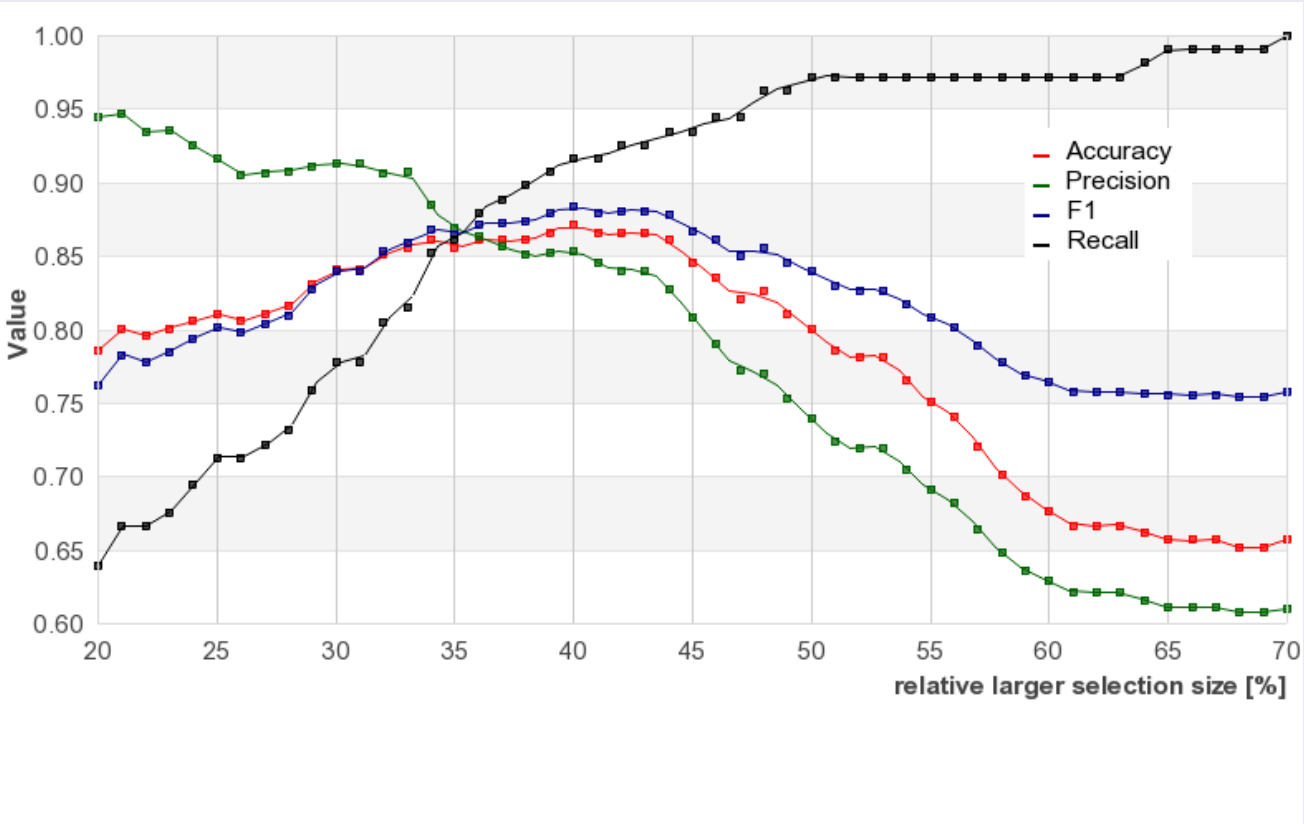


Figure: Measures of target size classifier in function of size metrics

Table of contents

1. Introduction
2. Framework for Describing Public Safety Video Applications
3. Classification of Video Sequences into Specified Generalized Use Classes
   - Research
   - Analysis of Results on Target Size
   - Analysis of Results on Lighting Level
4. Methods for Automatic Classification of Entire GUC Sequences
   - Data preparation
   - Target Size Classifier
   - Lighting Level Classifier
5. Method Evaluation
6. Conclusions and Further Development
7. Acknowledgements

Lighting Level

- VQiPS defining 3 levels of entire sequence illumination
  - Dim
  - Bright
  - Variable — rejected due to low stability
- Here, per-ROI responses also taken into account
- Finding binary classification criterion based on subjects

- Different numerical metrics of target sizes calculated
  - $F1$ — F1 score
  - A — Measuring accuracy
  - P — Precision
  - R — Recall
- $LL = avg(L_V(ROI))$
  - LL — Lighting Level metric
  - L_V — Luminance
- $A_{max}(LL = 55) \geq 80\%$

Lighting Level — Histogram

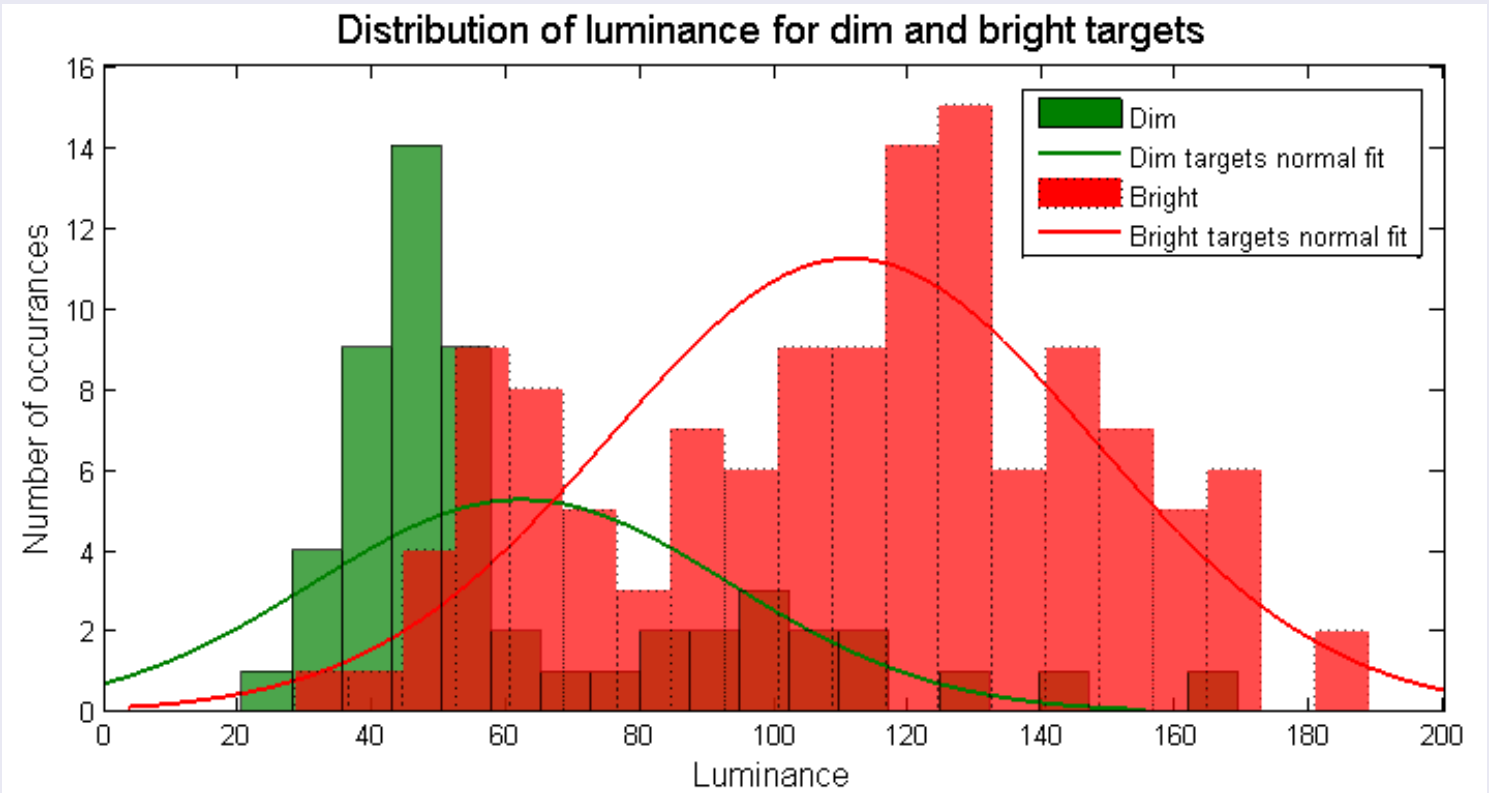


Figure: End-users selections of target lighting level in function of luminance

Lighting Level — Various Metric Values for Statistics

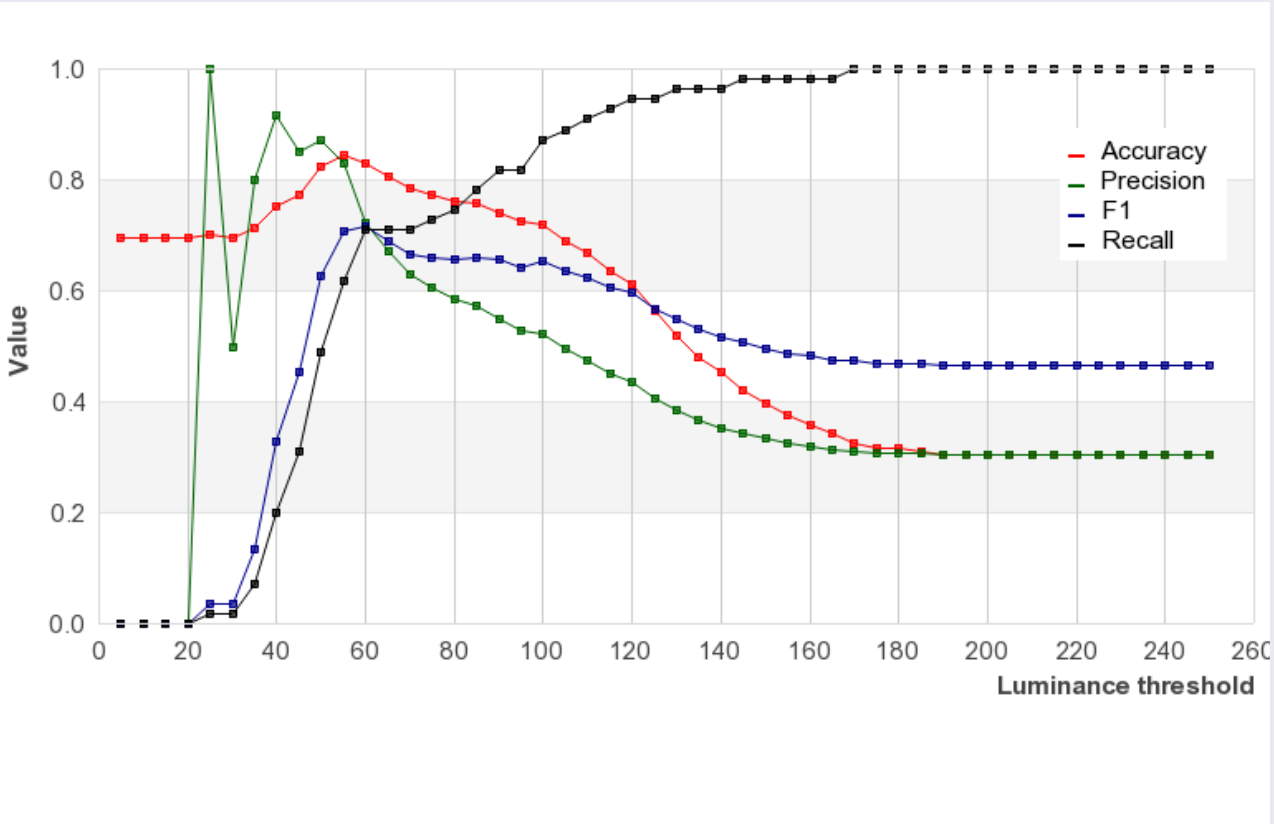


Figure: Measures of lighting level classifier in function of luminance threshold

Methods for Automatic Classification of Entire Generalized Use Class Sequences

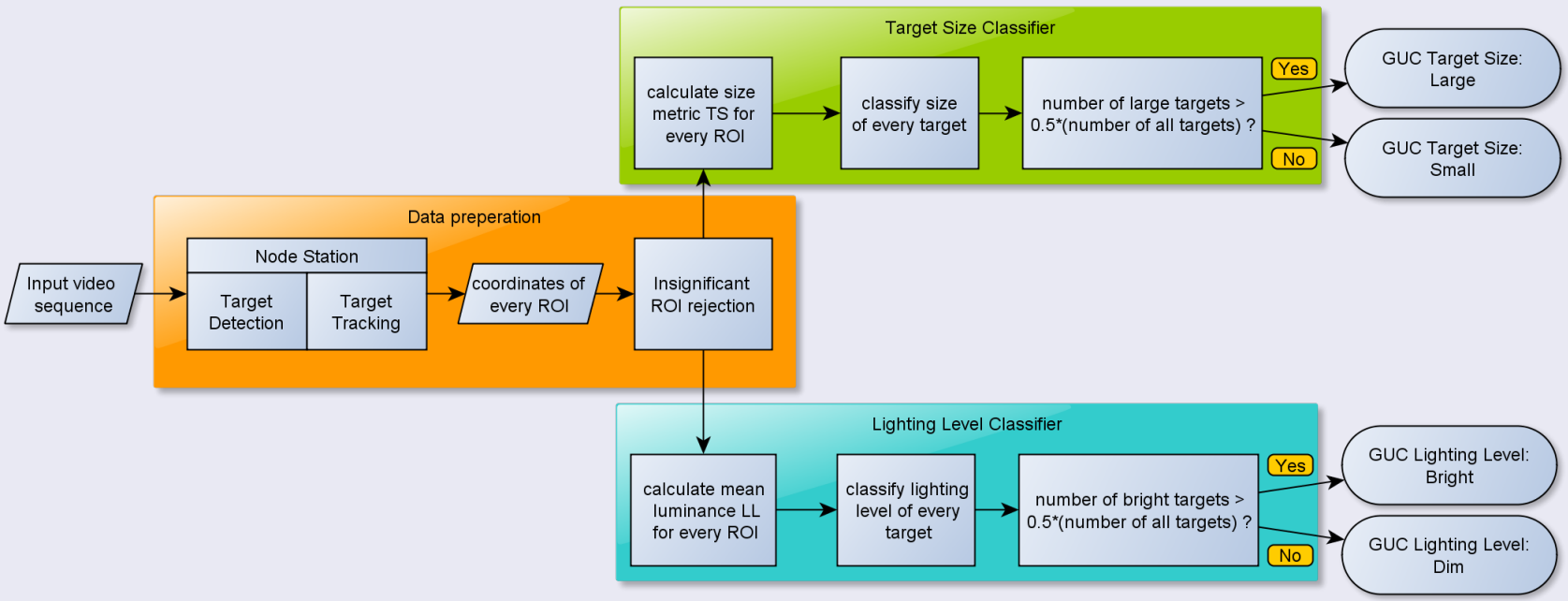
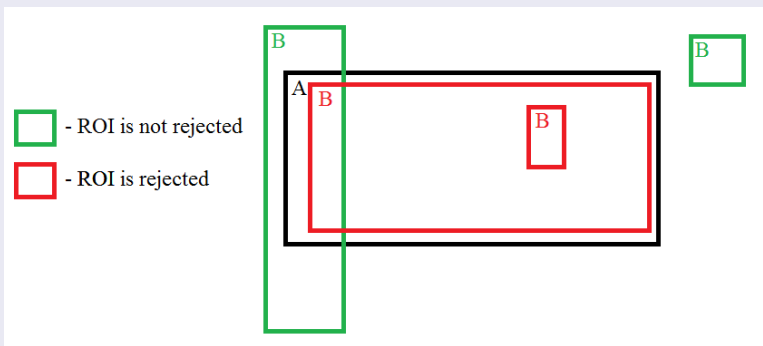


Figure: Block diagram of specified GUC classification method.

Table of contents

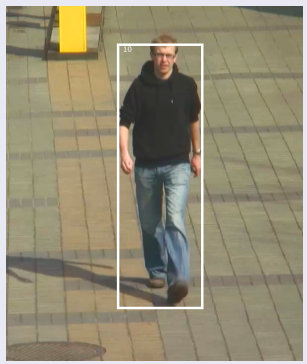
1. Introduction
2. Framework for Describing Public Safety Video Applications
3. Classification of Video Sequences into Specified Generalized Use Classes
   - Research
   - Analysis of Results on Target Size
   - Analysis of Results on Lighting Level
4. Methods for Automatic Classification of Entire GUC Sequences
   - Data preparation
   - Target Size Classifier
   - Lighting Level Classifier
5. Method Evaluation
6. Conclusions and Further Development
7. Acknowledgements

Data preparation



(a) Schema of rejection method

(b) Targets detected



(c) Targets following rejection

Figure: Reduction of detected redundant objects

Table of contents

1. Introduction
2. Framework for Describing Public Safety Video Applications
3. Classification of Video Sequences into Specified Generalized Use Classes
   - Research
   - Analysis of Results on Target Size
   - Analysis of Results on Lighting Level
4. Methods for Automatic Classification of Entire GUC Sequences
   - Data preparation
   - Target Size Classifier
   - Lighting Level Classifier
5. Method Evaluation
6. Conclusions and Further Development
7. Acknowledgements

Target Size Classifier

- Size metric of 40% used as threshold in binary classifier
- First task – to determine size metric for every significant ROI
- Calculated by dividing larger side of target selection by respective frame dimension:

$$TS = \frac{max(x, y)}{X \veebar Y} \qquad (1)$$

where:

- TS – "Target Size" metric
- x, y – size of selected ROI
- $X \veebar Y$ – respective length of frame dimension

- Every selection is classified as:
  - Large if $TS > 40\%$
  - Small if $TS \leq 40\%$
- The size of each target is obtained by a majority of sizes of the selection of the same target during the entire sequence
- After that, GUC Target Size parameter defined as majority of answers for all targets

Table of contents

1 Introduction
2 Framework for Describing Public Safety Video Applications
3 Classification of Video Sequences into Specified Generalized Use Classes
- Research
- Analysis of Results on Target Size
- Analysis of Results on Lighting Level

4 Methods for Automatic Classification of Entire GUC Sequences
- Data preparation
- Target Size Classifier
- Lighting Level Classifier

5 Method Evaluation
6 Conclusions and Further Development
7 Acknowledgements

Lighting Level Classifier

- The lighting level is selected by comparing the average luminance with the value of 55 – the threshold for which the highest accuracy occurs, as mentioned previously
- Classification starts with calculating of the mean luminance for every region of interest obtained in the data preparation step
- Firstly, the entire selection is converted into grey scale, and the mean luminance is calculated
- This value is compared to the value of 55 to determine lighting levels of each ROI
- Based on data from the tracker, lighting levels of each target are the same as the majority of lighting levels of its selections
- After that, the GUC lighting level parameter is defined as the majority of answers for all targets

Method Evaluation

- One moving group of pixels generally identified as 1 object
- But users recognized at 2+
- Therefore it was decided that:
  - Groups of moving objects are selected as a single target (for example, a group of running people)
  - Parts of targets moving together cannot be detected (for example, the face of a robber)
  - If two or more selection overlap, the larger one is taken into account
- Target size of entire sequence determined when $\frac{2}{3}$ of targets consistent with assumptions commonly determined by end-users
- Sequences randomly divided into:
  - Testing set
  - Training set
- Correlation with end-users opinions of:
  - 70% for object size
  - 93% for lighting level

Conclusions

- Size metric of 40% used as threshold in binary classifier of target size
- Lighting level selected by comparing average luminance with value of 55
- Subjects-driven methods for automatic classification of entire GUC sequence already developed
- Developed algorithms based on image processing of each video frame
- Target size classification with accuracy reaching 70% (satisfactory result indicating indecision of users)
- Lighting level classification with accuracy reaching 93%
- Computer classification of any footage into GUCs cannot be taken as certain result, therefore it should be verified manually

Further Development

- Main issue that emerged during evaluation of automatic methods of classification into specified GUCs was imperfection of detector
- Development of this module by implementation of following methods will significantly improve range of applications for system:
  - Detection of sub-objects (such as a weapon)
  - Detection of stationary objects (such as abandoned luggage)
  - Detection of targets at sequences containing moving background (such as footage recorded in car during pursuit)
- This research to be also contribution to study on automatic classification of motion level
- Planned combination with NTIA research on motion level

Acknowledgements

The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under Grant Agreement №218086 (INDECT).