Illumination Robustness of Moving Object Identification

Using Interval Spatio-Temporal Graphs of Parameterized Contours

Andrii Blyndaruk · Olena Shapovalova

Simon Kuznets Kharkiv National University of Economics, Kharkiv, Ukraine

Keywords: Computer Vision · Moving Object Identification · Parameterized Contours · NURBS · Interval Spatio-Temporal Graphs · Neural Networks · Illumination Robustness · Structural State Graphs (SSG)

Motivation and Aim

Illumination variation directly affects the pixel-level representation of a scene.

Problem context

Moving object identification is used in video surveillance, robotics, unmanned and transportation systems, and automated monitoring.

Practical conditions may involve changes in background, partial occlusions, noise, and variations in illumination.

Changes in illumination may affect segmentation and contour extraction, and consequently the subsequent geometric and structural representation.

Aim of the study

To experimentally investigate the robustness of the proposed method for modeling and identifying moving object behavior to reduced illumination in test video data without retraining the neural network model.

The study evaluates identification results at the level of interval spatio-temporal graphs and complete video recordings.

Proposed Method: From Video to Video-Level Identification

Complete computational pipeline of the proposed method.

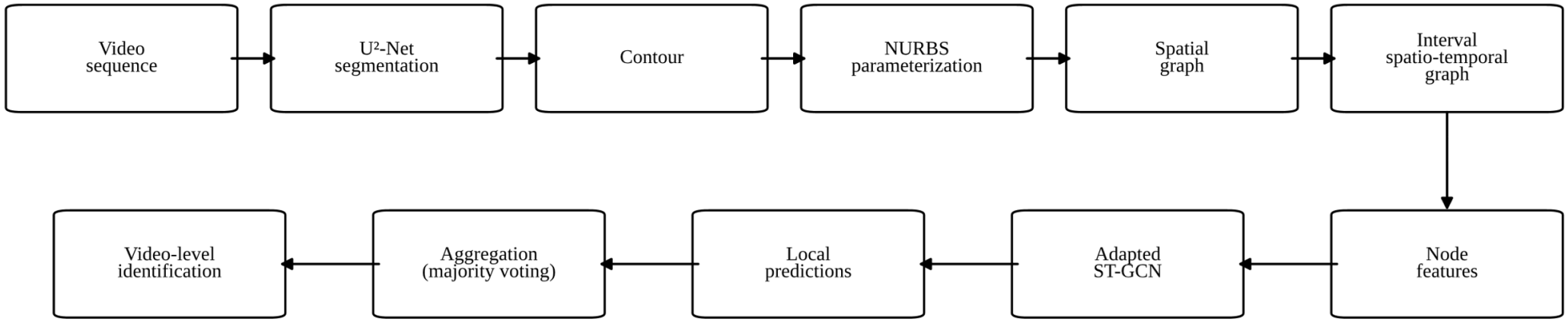


Fig. 2. General pipeline of the proposed method for moving object behavior identification.

video sequence → U^2-Net segmentation → contour → NURBS parameterization → spatial graph → interval spatio-temporal graph → node features → adapted ST-GCN → local predictions → aggregation → video-level identification

Structured Representation: Interval Spatio-Temporal Graph

The primary structured unit of analysis in the proposed method.

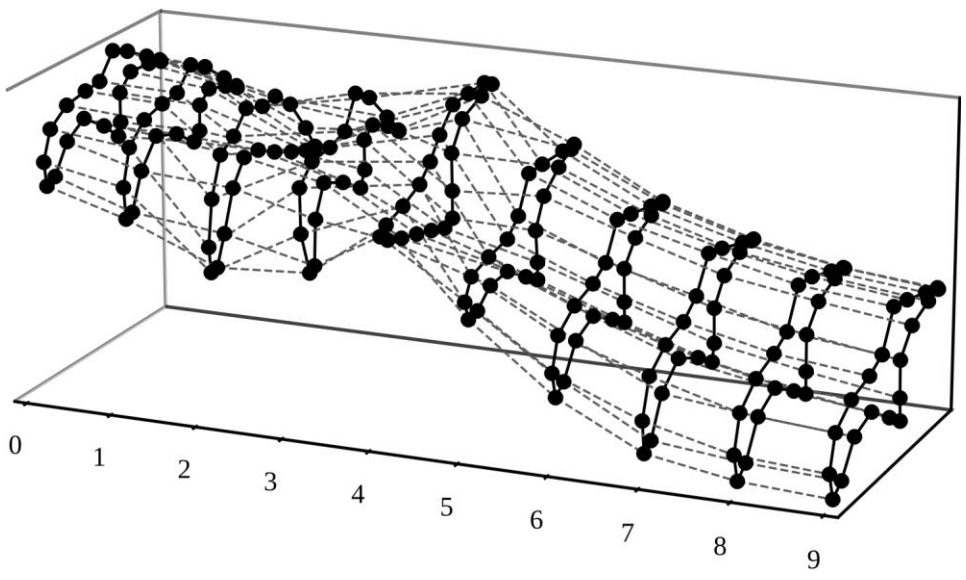


Fig. 1. Interval spatio-temporal graph for 10 consecutive frames.

Graph construction

Spatial graph with 20 nodes is constructed for each valid frame.

Sliding window of 10 frames with a stride of 1 is used.

Each interval graph contains 200 nodes, 200 spatial edges, and 180 temporal edges.

Each node is characterized by 13 geometric, positional, and dynamic features.

SSG denotes a graph-based representation of object states and relationships between them.

Experimental Setup: Cross-Light Evaluation

The model was trained once and applied to darkened test videos without adaptation.

Dataset	Split	Graphs	Model
15 videos 3 classes	9 / 3 / 3 Train / Val / Test	760 interval graphs	5,507 parameters

Fixed experimental configuration

Frames were sampled with a step of 3, with no more than 60 frames processed from each video.

NURBS: degree = 3; 20 discretized points.

Interval graph: 10 frames; stride = 1; 13 node features.

Illumination protocol

Training and validation data remained unchanged with respect to illumination.

Brightness modification was applied only to the test video data.

Brightness coefficients: α = 1.00, 0.75, 0.60, 0.55, 0.50, 0.25, 0.10, 0.05, 0.005.

No retraining · No fine-tuning · Same fixed model state

Interval-Level Results under Reduced Illumination

Accuracy did not exhibit a monotonic decrease as brightness was reduced.

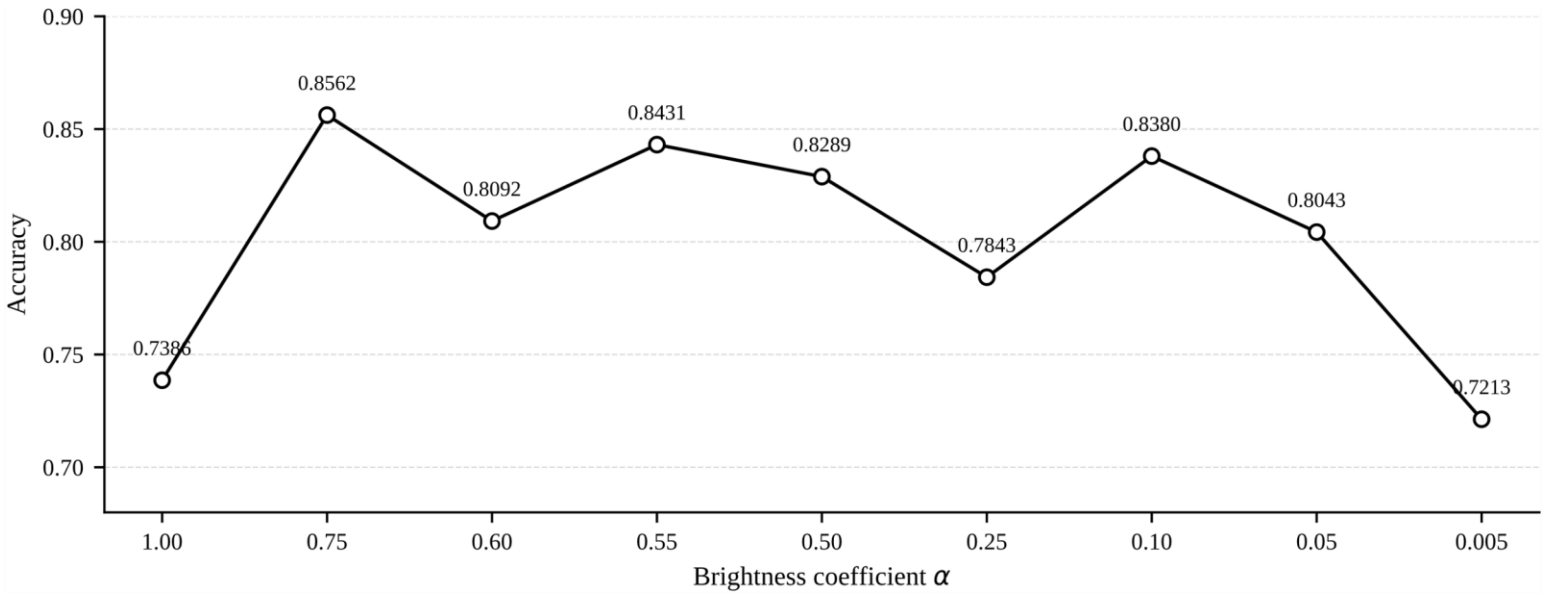


Fig. 3. Identification accuracy at the interval spatio-temporal graph level.

Key results

Accuracy ranged from 0.7213 to 0.8562 across the entire investigated brightness range.

Maximum Accuracy of 0.8562 was obtained at α = 0.75.

At α = 0.005, Accuracy was 0.7213 for successfully formed interval graphs.

Video-Level Identification: Main Practical Outcome

All generated final predictions were correct at every investigated brightness level.

α = 1.00–0.05

3 / 3 correct at each level

Coverage

1.0000 for α = 1.00–0.05

α = 0.005

2 / 2 correct coverage 0.6667

The relationship between interval-level and video-level results is informative: individual incorrect local predictions do not determine the identification result for the entire video when the majority of intervals correctly capture characteristic structural and dynamic properties.

For the third video at α = 0.005, no prediction was generated because a sufficient sequence of valid contours could not be formed.

This case does not represent a misclassification by the adapted ST-GCN.

Conclusions and Future Work

Structured spatio-temporal representation preserved final identification under changing illumination.

Conclusions

The cross-light study quantitatively confirmed the effectiveness of the proposed method.

Final identification remained correct for every test video for which the required structured graph representation was formed.

The results support the practical potential of structured spatio-temporal representation under changing illumination conditions.

Future work

Expanding the experimental dataset.

Evaluating the method on a larger number of classes and independent video recordings.

Investigating natural and nonuniform illumination changes.

Further developing adaptive formation of temporal graph representations under challenging observation conditions.

Thank you