LIFE-LONG VISUAL LOCALIZATION USING PROBABILISTIC TEMPORAL INFERENCE

Arroyo R., Bergasa L. M., Romera E. - University of Alcalá (UAH)
{roberto.arroyo, bergasa, eduardo.romera}@depeca.uah.es

Alcantarilla P. F. - Toshiba Research Europe Ltd.
pablo.alcantarilla@crl.toshiba.co.uk

ICVSS 2015
Sicily ~ 12-18 July
International Computer Vision Summer School

Universidad de Alcalá

TOSHIBA
Leading Innovation >>>

ABSTRACT

Life-long visual localization is a challenging computer vision topic because of the strong appearance transformations that a place usually suffers due to geometric changes, dynamic elements, weather or seasons. We have designed a method named ABLE which can visually recognize locations at different times of day, along the months or seasons. Our current objective is also the application of a probabilistic detection of geometric changes across the four seasons in 3D reconstructed environments.

MOTIVATION AND OVERVIEW

The main goal of our research is to re-identify places using visual information and to detect changes in these locations, which are critical tasks in any visual localization system. Besides, a temporal inference based on probabilities can help to reduce the uncertainty associated with **large-scale problems**.

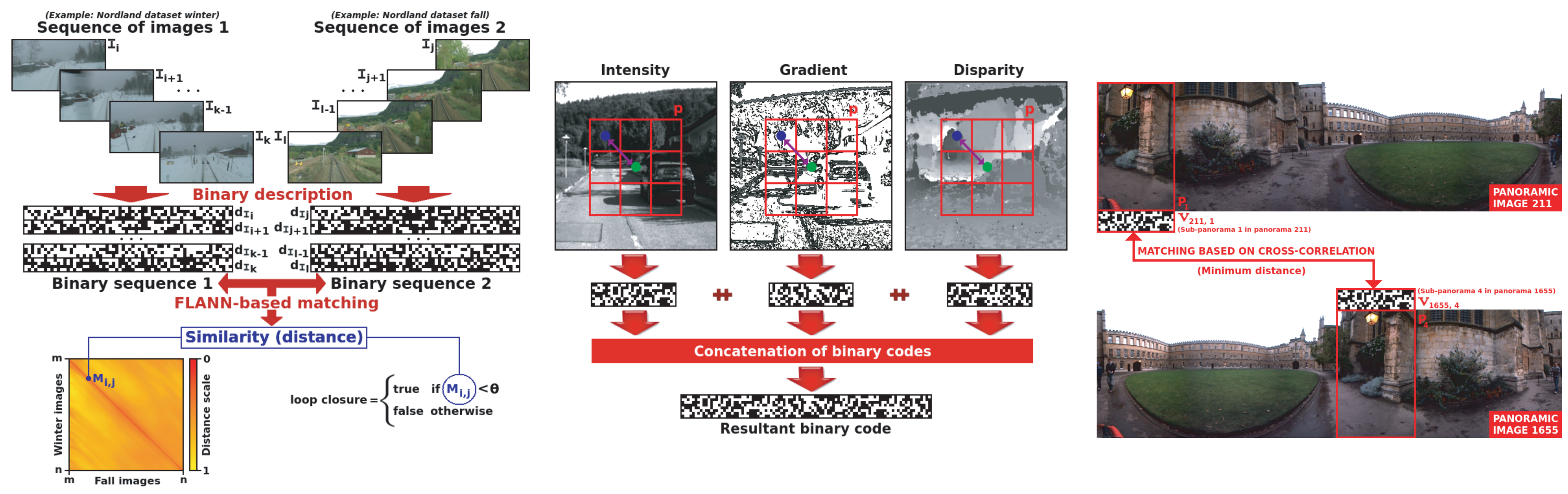
Nowadays, state-of-the-art works are starting to study the difficulties of **visual localization in a long-term context**. As the most representative example, **seasonal changes** produce drastic modifications in the visual appearance of places.

a) Spring b) Summer c) Fall d) Winter

Fig. 1. An example of seasonal changes in a same place in the Nordland dataset.
The Nordland dataset is available in:
https://nrkbeta.no/2013/01/15/nordlandsbanen-minute-by-minute-season-by-season/

OUR METHOD FOR VISUAL PLACE RECOGNITION: ABLE

ABLE (**A**ble for **B**inary-appearance **L**oop-closure **E**valuation) is our proposal for re-identifying places using only visual appearance. Images are described as binary codes extracted from a global **LDB** (Local Difference Binary) descriptor and efficiently matched using FLANN. This approach provides a **great reduction of memory and computational costs**, which is necessary for long-term performance.

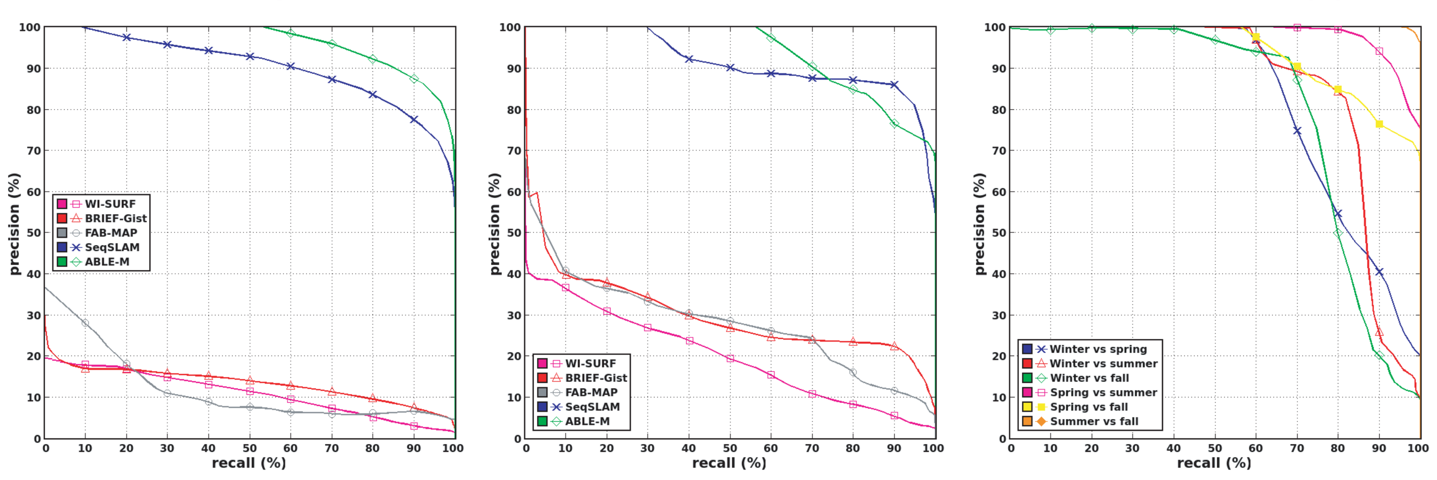


a) ABLE-M [1] (Monocular images) b) ABLE-S [2] (Stereo images) c) ABLE-P [3] (Panoramic images)

Fig. 2. The different versions of ABLE. Check out our published material in [1] [2] [3] for a deeper explanation of each approach.

EXPERIMENTAL RESULTS IN VISUAL PLACE RECOGNITION FOR ABLE-M

Our method for place recognition is satisfactorily compared against the state of the art. Evaluation is based on similarity matrices and precision-recall curves. We show results for our last approach: **ABLE-M**. For more tests, see [1] [2] [3].



a) ABLE-M vs state-of-the-art methods (CMU-CVG Visual Localization dataset) b) ABLE-M vs state-of-the-art methods (Nordland dataset) c) ABLE-M performance between seasons (Nordland dataset)

Fig. 3. P-R curves showing the performance of ABLE-M in challenging life-long localization scenarios.

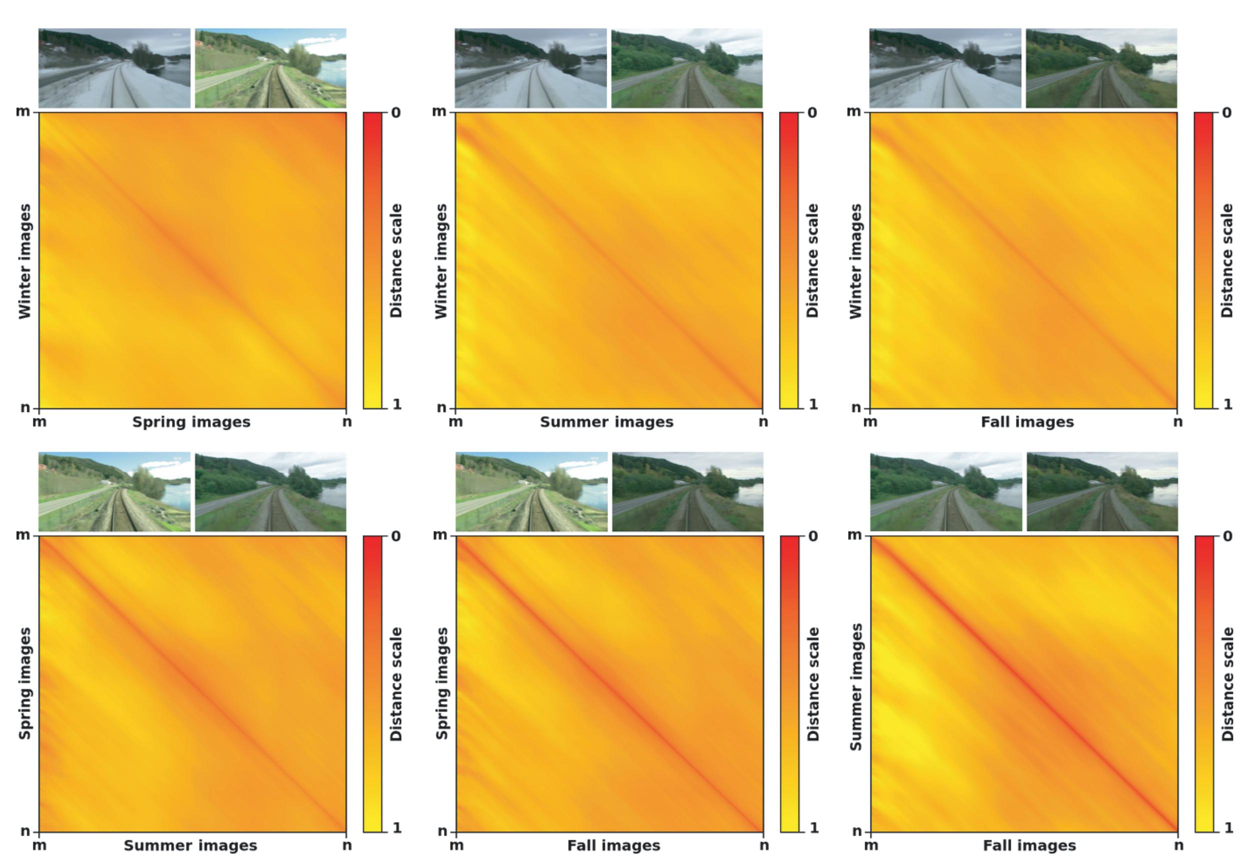


Fig. 4. Similarity matrices showing ABLE-M performance across the four seasons in the Nordland dataset.

Why ABLE-M improves state-of-the-art methods in life-long visual localization?

- **Sequences of images** instead of single images for recognizing places.

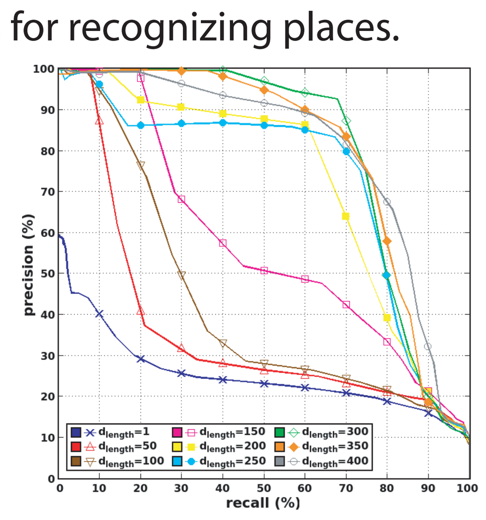
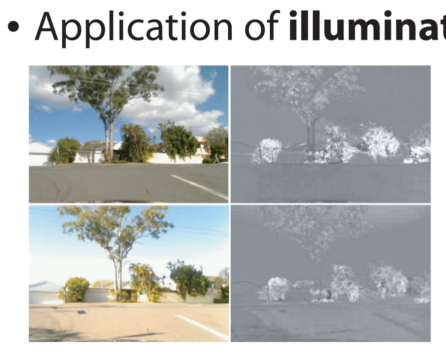


Fig. 5. ABLE-M performance depending on the length of the sequences of images (Nordland dataset).

- Application of **illumination invariance**.



$$\mathcal{I} = \log(G) - \alpha \cdot \log(B) - (1-\alpha) \cdot \log(R) \quad (1)$$

$$\frac{1}{\lambda_G} = \frac{\alpha}{\lambda_B} + \frac{(1-\alpha)}{\lambda_R} \quad (2)$$

$$\alpha = \frac{\left(\frac{\lambda_R}{\lambda_G} - \frac{\lambda_R}{\lambda_B}\right)}{\left(1 - \frac{\lambda_R}{\lambda_B}\right)} \quad (3)$$

Fig. 6. Example of illumination invariance usage and basic formulation for RGB invariant transformation.

- The **binary descriptors** applied by ABLE-M jointly with the matching based on **Hamming distance and FLANN** provide a remarkable performance and low processing times.

FUTURE WORK: CHANGE DETECTION

The **detection of geometric changes** in places traversed on different months of the year is the future goal of our research.

3D reconstructions of the environment are carried out in different periods of time. After that, a grid mapping comparison based on probabilistic temporal inference is performed with the aim of evaluating the geometric changes that a place has suffered along the time.

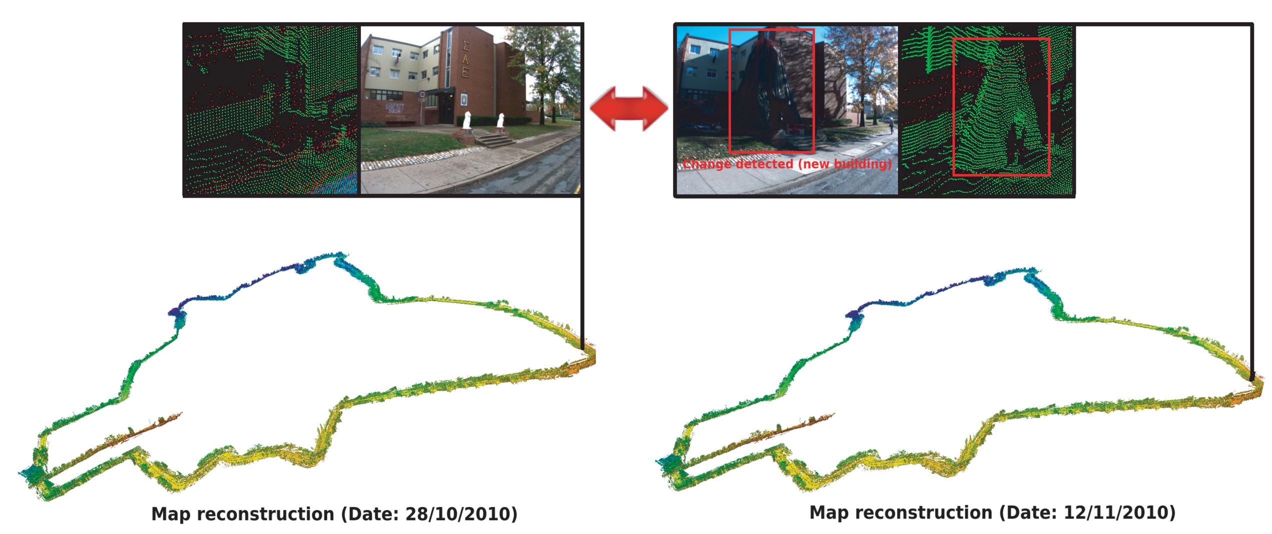


Fig. 7. An example of geometric change detection in the CMU-CVG Visual Localization dataset.
The CMU-CVG Visual Localization dataset is available in:
http://3dvis.ri.cmu.edu/data-sets/localization/

CONCLUSIONS

Our work contributes an innovative approach for visual localization, which has been successfully tested in life-long scenarios. Besides, ABLE is an efficient alternative to state-of-the-art methods.

For more information about our work, please visit this Webpage:
http://www.robesafe.uah.es/personal/roberto.arroyo/

REFERENCES

[1] Arroyo R., Alcantarilla P. F., Bergasa L. M., Romera E., "Towards Life-Long Visual Localization using an Efficient Matching of Binary Sequences from Images", in *IEEE International Conference on Robotics and Automation (ICRA)*, Seattle, Washington (USA), May 2015.

[2] Arroyo R., Alcantarilla P. F., Bergasa L. M., Yebes J. J., Bronte S., "Fast and effective visual place recognition using binary codes and disparity information," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Chicago, Illinois (USA), September 2014, pp. 3089–3094.

[3] Arroyo R., Alcantarilla P. F., Bergasa L. M., Yebes J. J., Gámez S., "Bidirectional loop closure detection on panoramas for visual navigation," in *IEEE Intelligent Vehicles Symposium (IV)*, Dearborn, Michigan (USA), June 2014, pp. 1378–1383.

ACKNOWLEDGEMENTS

This work is funded by the UAH through a FPI grant, the Spanish MINECO through the project Smart Driving Applications (TEC2012-37104) and the CAM through the project RoboCity2030-III-CM (P2013/MiT2748).

Special thanks to the people of Toshiba Research Europe for their support during my internship in the Cambridge Research Lab. My thanks also goes to all my colleagues in the RobeSafe research group for their constant help.