Application

2015-04-02

2015-05614	Flierl, Markus	NT-14

Information about applicant

Name: Markus Flierl

Doctorial degree: 2003-06-27

Birthdate: 19710704

Academic title: Docent

Gender: Male

Employer: Kungliga Tekniska högskolan

Administrating organisation: Kungliga Tekniska högskolan

Project site: Avdelningen för Kommunikationsteori

Information about application

Call name: Forskningsbidrag Stora utlysningen 2015 (Naturvetenskap och teknikvetenskap)

Type of grant: Projektbidrag

Focus: Fri

Subject area:

Project title (english): Mobile 3D Visual Search

Project start: 2016-01-01

Project end: 2019-12-31

Review panel applied for: NT-14

Classification code: 20204. Telekommunikation

Keywords: Mobile visual search, data compression, image-based features, object recognition

Funds applied for

Year:	2016	2017	2018	2019
Amount:	1,496,000	1,544,200	1,699,100	1,779,400

Descriptive data

Project info

Project title (Swedish)*

Mobile 3D Visual Search

Project title (English)*

Mobile 3D Visual Search

Abstract (English)*

Visual information is important in our daily lives and we aim to use it efficiently. The advancement of wireless mobile devices allows us to immerse ourselves in a new augmented reality. We consider the fact that the human being has binocular vision
and is mobile. This allows us to consider three-dimensional real-world objects, collect visual and geometric (shape) information, and transmit extracted data to remote data centers for further processing.

This project focuses on the search and retrieval of three-dimensional real-world objects that a mobile user may accomplish by sending a query over a communication link to a remote data center and retrieving the desired database information. The query
includes both visual and geometric information to identify real-world objects uniquely. A central aspect is the relation between the capacity of the communication link and the quality of the retrieved information. Further, an efficient representation of a large
number of three-dimensional objects in a database is desirable for optimal storage and retrieval.

The goal of this project is to develop mathematical models and algorithms for efficient communication-constrained three-dimensional object search. We expect that search based on visual and geometric information will be advantageous to uniquely identify buildings, structures, and unknown objects on the road in autonomous driving situations.
Further, the communication constraint will help to identify the features that are relevant for three-dimensional object search. Due to limited communication resources, only the relevant features will be transmitted and used for search.

Popular scientific description (Swedish)*

Visual information is important in our daily lives and we aim to use it efficiently. The advancement of wireless mobile devices allows us to immerse ourselves in a new augmented reality. We consider the fact that the human being has binocular vision
and is mobile. This allows us to consider three-dimensional real-world objects, collect visual and geometric (shape) information, and transmit extracted data to remote data centers for further processing.

This project focuses on the search and retrieval of three-dimensional real-world objects that a mobile user may accomplish by sending a query over a communication link to a remote data center and retrieving the desired database information. The query
includes both visual and geometric information to identify real-world objects uniquely. A central aspect is the relation between the capacity of the communication link and the quality of the retrieved information. Further, an efficient representation of a large
number of three-dimensional objects in a database is desirable for optimal storage and retrieval.

The goal of this project is to develop mathematical models and algorithms for efficient communication-constrained three-dimensional object search. We expect that search based on visual and geometric information will be advantageous to uniquely identify buildings, structures, and unknown objects on the road in autonomous driving situations.
Further, the communication constraint will help to identify the features that are relevant for three-dimensional object search. Due to limited communication resources, only the relevant features will be transmitted and used for search.

Project period

Number of project years*

4

Calculated project time*

2016-01-01 - 2019-12-31

Classifications

Select a minimum of one and a maximum of three SCB-codes in order of priority.

Select the SCB-code in three levels and then click the lower plus-button to save your selection.

SCB-codes*	2. Teknik > 202. Elektroteknik och elektronik > 20204. Telekommunikation

---

Enter a minimum of three, and up to five, short keywords that describe your project.

Keyword 1*

Mobile visual search

Keyword 2*

data compression

Keyword 3*

image-based features

Keyword 4

object recognition

Keyword 5

Ethical considerations

Specify any ethical issues that the project (or equivalent) raises, and describe how they will be addressed in your research. Also indicate the specific considerations that might be relevant to your application.

Reporting of ethical considerations*

no ethical considerations.

The project includes handling of personal data

No

The project includes animal experiments

No

Account of experiments on humans

No

Research plan

Mobile 3D Visual Search

Visual information is important in our daily lives and we aim to use it efficiently. The advancement of wireless mobile devices allows us to immerse ourselves in a new augmented reality. We consider the fact that the human being has binocular vision and is mobile. This allows us to consider three-dimensional real-world objects, collect visual and geometric (shape) information, and transmit extracted data to remote data centers for further processing.

This project focuses on the search and retrieval of three-dimensional real-world objects that a mobile user may accomplish by sending a query over a communication link to a remote data center and retrieving the desired database information. The query includes both visual and geometric information to identify real-world objects uniquely. A central aspect is the relation between the capacity of the communication link and the quality of the retrieved information. Further, an efficient representation of a large number of three-dimensional objects in a database is desirable for optimal storage and retrieval.

The goal of this project is to develop mathematical models and algorithms for efficient communication-constrained three-dimensional object search. We expect that search based on visual and geometric information will be advantageous to uniquely identify buildings, structures, and unknown objects on the road in autonomous driving situations. Further, the communication constraint will help to identify the features that are relevant for three-dimensional object search. Due to limited communication resources, only the relevant features will be transmitted and used for search.

1 State-of-the-Art and Objectives

This project focuses on the search and retrieval of three-dimensional (3D) real-world objects that a mobile user may accomplish by sending a query over a communication link to a remote database and retrieving the desired database information.

1.1 Visual Search

The authors of "Video Google" [1] used a text retrieval approach to object matching in videos. First, image regions have been represented by 128-dimensional vectors using the Scale Invariant Feature Transform (SIFT) descriptor [2]. Second, the descriptors have been vector quantized into clusters which are the visual words for text retrieval. This way, a visual vocabulary can be constructed from a subset of images. Finally, the visual vocabulary has been used to retrieve objects in the remaining images of the video. In this early work of visual search, no consideration was given to the volume of data consumed by the descriptor vectors.

1.2 Mobile Visual Search

The advancement of wireless mobile devices and the desire for an augmented reality in a real-world environment have raised the interest in mobile visual search [3]. With the integration of digital cameras into mobile devices, image-based information retrieval for mobile visual search has been developing rapidly. A crucial problem is the efficient utilization of the information in the mobile images. The challenges of mobile image retrieval are rooted in the bandwidth constraint and the limited computational capacity of mobile devices. To solve these problems, most of the mobile visual search algorithms use the so-called bag of features approach where only the salient image features are extracted and sent. Therefore, the detection of features and the computation of descriptors play an important role in feature extraction. The SIFT descriptor [2] has been widely used in visual search applications. It is more robust than many other well-known features in the context of feature matching and recognition due to its invariance under rotation, scale change and affine transformation [4].

However, the direct transmission of SIFT descriptors is not practical due to the large data volume. In particular, the amount of SIFT data is usually larger than the size of the JPEGcompressed image itself [3]. Hence, several compression schemes have been proposed to solve this problem. An efficient approach is known as Compressed Histogram of Gradients (CHoG) [5], which can reduce the data rate by factor 20 when compared to that of the uncompressed SIFT descriptor. However at very low data rate, the recall rate decreases significantly as only a few features can be matched correctly.

1.3 Theoretical Models for Similarity Queries

As mentioned above, visual words are vector quantized descriptors and the resulting visual vocabulary is a compressed version of the 3D objects in the database. With that, the problem of communication-constrained object search turns into a problem of performing similarity queries on compressed data. The work in [6] studies the fundamental trade-off between compression rate, "sequence" length, and reliability of queries performed on compressed data. A so-called identification rate is defined for an abstract query "sequence" of given length. When compression is performed at a rate greater than the identification rate, responses to queries on the compressed data can be made exponentially reliable. The work focuses on the quadratic distance between "sequences" for the similarity criterion. Two compression schemes for similarity queries have been developed in a follow-up work [7].

These models consider only abstract query "sequences". For communication-constrained object search, lossy compression of query "sequences" is necessary in order to achieve short description lengths for efficient communication.

1.4 Objectives

The main objective of this project is to develop mathematical models and algorithms to accomplish efficient communication-constrained 3D object search. This efficient search is characterized by the following aspects:

- The queries include both visual and geometric (shape) information to identify 3D objects uniquely. If the visual information of two 3D objects is very similar, the geometric information can be used to discriminate them. For example, we like to distinguish between a given real-world building and a flat poster showing that building. Obviously, the visual information alone will not be helpful, and the geometric information will be necessary.

- A minimum description length of the query (number of bits to be sent over the communication link) is necessary to identify the 3D object in the database uniquely. It is assumed that the minimum description length depends on the number of similar 3D objects in the database. To our understanding, there is a fundamental trade-off between the description length and the probability of correct recognition. This trade-off suggests an embedded description for the queries to be sent over the communication link.

- A compressed representation for a large number of 3D objects in the database is necessary for efficient storage. The representation for the 3D objects needs to consider both the visual and the geometric information. For visual information, so-called vocabulary trees have been used. They permit also efficient compression. However, the integration of geometric information is not obvious.

- A minimum size of the compressed representation is necessary to identify each 3D object in the database uniquely. Here, theoretical models for similarity queries may be helpful to determine the so-called identification rate. As mentioned above, compression at a rate greater than the identification rate makes responses to queries exponentially reliable.

We see the relevance of this project as follows:

- Through information technology, more and more data becomes available, and each user faces an increasing challenge to find relevant information by intelligent "search".
- Today, search is not limited to text retrieval. Visual search based on visual words is used by Firefly, a feature of the Amazon Fire Phone [8]. Visual search captures only the visual appearance of objects and uses it for retrieval.
- The visual appearance of 3D objects is not always sufficient to discriminate between two 3D objects. Hence, it is necessary to capture also geometric (shape) information to identify 3D objects uniquely.
- Search based on visual and geometric information may be helpful to identify buildings, structures, and unknown objects on the road in autonomous driving situations.
- In contrast to short text queries, such 3D object queries may burden a (wireless) communication link. Hence, communication-efficient queries are desirable.

This list sketches the scientific relevance of the project and points out an opportunity for new technologies.

1.5 Impact

This project is in response to the growing need to utilize visual information in our daily lives. We consider the fact that the human being has binocular vision and is mobile. For example, search based on visual and geometric information will be advantageous to uniquely identify buildings, structures, and unknown objects on the road in autonomous driving situations. Further, the communication constraint will help to identify the features that are relevant for 3D object search. Due to limited communication resources, only the relevant features will be transmitted and used for search.

2 Methodology

Image-based features like SIFT are well suited to capture the visual appearance of objects in an image. However, with one view only, no geometric information can be extracted. Therefore, we use for the same object a set of multi-view images to define so-called multi-view features. For each image, we extract a set of image-based features. SIFT features are preferred due to their robustness under rotation, scale change and affine transformation [2]. However, as there are multiple feature sets for the same object, the redundancy in feature space is high. Thus, efficient feature selection algorithms are needed to discriminate features. On the other hand, as the images are taken from different perspectives with varying lighting conditions, using a robust representation for the feature descriptors is important for reliable feature matching.

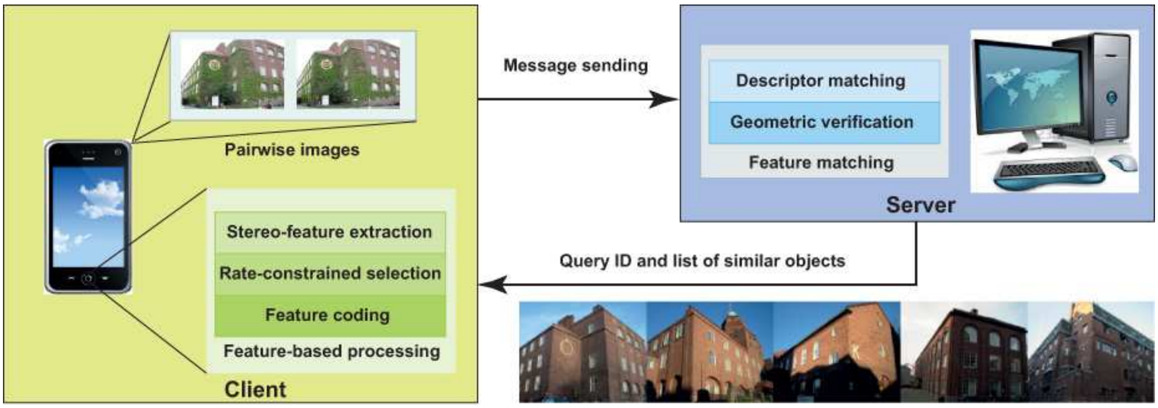

Figure 1: Query process for communication-constrained 3D object search.

Fig. 1 depicts an example query process for communication-constrained 3D object search to highlight the various aspects of the project.

2.1 Rate-Constrained Feature Selection

For each query, we transmit a set of multi-view features over a communication link with given bandwidth. In order to meet the bandwidth constraint, we sample the acquired multi-view features and send only the most reliable candidates. The most reliable candidates are those that increase the probability of correct recognition maximally. Hence, we will work on algorithms that identify the most reliable candidates. An example is given in our prior work [9].

2.2 Hierarchically Structured Multi-View Features

Multi-view feature correspondences provide a way to measure the reliability of features. A simple feature correspondence relates a feature point in image A to a feature point in image B. A multi-view feature correspondence in l images relates l feature points, where each image contributes exactly one feature point. That is, a multi-view feature correspondence is represented by a undirected complete graph with l vertices. The i-th vertex in the graph represents a feature in the i-th image. Each edge represents the correspondence between two features. A complete graph with more vertices is more reliable and representative than a complete graph with less vertices. Simply speaking, a large complete graph represents correspondences among many images.

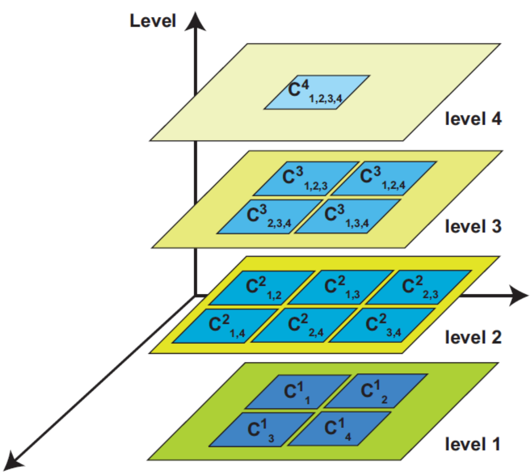

Figure 2: Hierarchical sets of features with four levels from four views.

As multi-view features with more vertices are more reliable for robust matching, we structure the sets of feature correspondences in a hierarchical manner, as shown in **Fig. 2**. In particular, sets of feature correspondences with l vertices are placed on level l. Further, subsets with $k < l$ vertices, where feature correspondences can be established, are placed on level k. Hence, we will work on algorithms that identify multi-view features. An example is given in our prior work [10].

Note, the hierarchically structured multi-view features should not be confused with well-known vocabulary trees [11]. Both are complementary structures. With hierarchically structured multi-view features, we are able to generate vocabulary trees which are based on multi-view feature descriptors. Both hierarchical and tree structures provide us an advantage when constructing representations for 3D objects.

As discussed above, we represent multi-view feature correspondences by undirected complete graphs with l vertices. However, this leads to l similar feature descriptor vectors that need to be represented efficiently. Our related work on compression of data on graphs [12] may be helpful for this problem.

2.3 3D Geometric Verification

To obtain geometrically correct matches for our queries, we need to introduce geometric verification into the matching process. Usually, an epipolar-constrained RANSAC algorithm [13] is used for single-view visual search. However, the epipolar constraint [14] can not capture the 3D information of the object. With our multi-view approach, the 3D geometric information of multi-view features can be efficiently used for geometric verification.

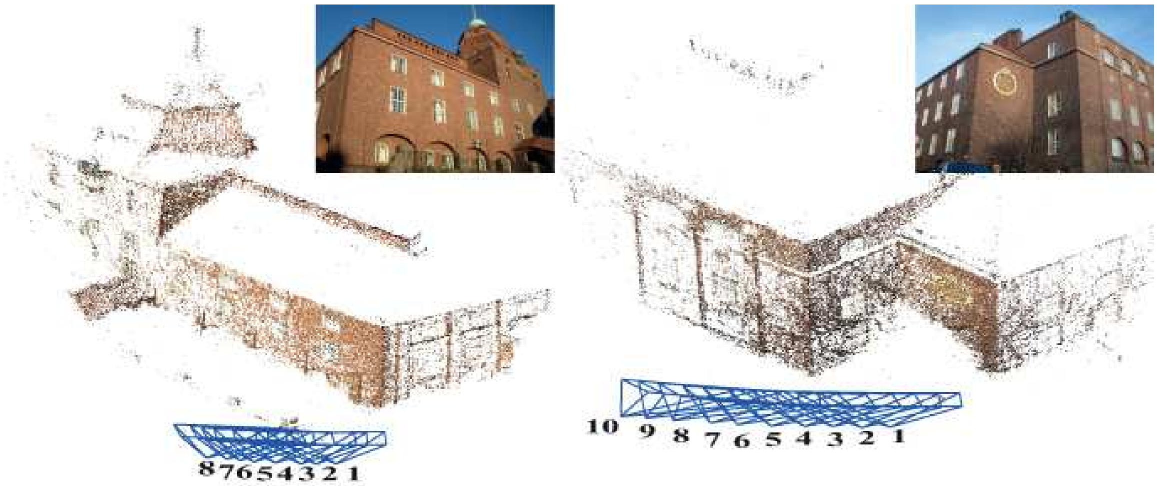

Figure 3: 3D geometric information as extracted from multi-view images.

Fig. 3 depicts an example for 3D geometric information as extracted from multi-view images. We will work on algorithms that use such information in our communication-constrained setting to identify 3D objects uniquely. An example is given in our current work [15].

2.4 Models for Communication-Constrained Object Search

Our goal is to develop a mathematical framework for communication-constrained object search. As discussed above, lossy compression of query "sequences" is necessary in order to achieve short description lengths for efficient communication. Further, a similarity criterion needs to be established that captures both visual and geometric similarity of 3D objects.

3 Resources and Project Costs

3.1 Research Team

The PI will implement this project in the context of his research group at KTH Royal Institute of Technology. The team that will work on the project will include one Ph.D. students with background in signal processing, coding, and computer vision. The student will focus on the relation between the capacity of the communication link and the quality of the retrieved information. He/she will also study the problem of efficient representations for a large number of three-dimensional objects in a database.

The PI will spend 40% of his working time to lead the project and to supervise the student throughout the duration of the project.

3.2 Other Resources

The project will be implemented based on the infrastructure provided by the Communication Theory Laboratory at the School of Electrical Engineering. Students and postdocs will be equipped with laptops and external monitors. In addition, a stand-alone workstation will be set up for running software simulations. Additional simulation capacity is available via the infrastructure provided by the School of Electrical Engineering.

3.3 Project Costs

During the four-year period at KTH, the annual direct costs for a Ph.D. student are expected to be about 800k SEK. The PI will invest about 40% of his time in the project. This includes the supervision of the student. The corresponding annual direct costs for the PI are expected to be about 450k SEK on average.

Equipment costs will include laptops with external monitors and one stand-alone workstation for simulation purposes. The total costs for equipment are estimated at 120k SEK.

Travel costs are included to present research findings at international conferences. Indirect costs include rental expenses for office space, IT support, and administrative support.

References

[1] J. Sivic and A. Zisserman, “Video Google: A text retrieval approach to object matching in videos”, in *Proceedings of the International Conference on Computer Vision*, Nice, France, Oct. 2003.

[2] D. Lowe, “Distinctive image features from scale-invariant keypoints”, *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.

[3] B. Girod, V. Chandrasekhar, R. Grzeszczuk, and Y. Reznik, “Mobile visual search: Architectures, technologies, and the emerging MPEG standard”, *IEEE Transactions on Multimedia*, vol. 18, no. 3, pp. 86–94, Mar. 2011.

[4] K. Mikolajczyk and C. Schmid, “A performance evaluation of local descriptors”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 10, pp. 1615–1630, 2005.

[5] V. Chandrasekhar, G. Takacs, D. Chen, S. Tsai, R. Grzeszczuk, and B. Girod, "CHoG: Compressed histogram of gradients – A low bit-rate feature descriptor", in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Miami Beach, FL, June 2009.

[6] A. Ingber, T. Courtade, and T. Weissman, "Quadratic similarity queries on compressed data", in *Proceedings of the Data Compression Conference*, Snowbird, UT, Mar. 2013.

[7] I. Ochoa, A. Ingber, and T. Weissman, "Compression schemes for similarity queries", in *Proceedings of the Data Compression Conference*, Snowbird, UT, Mar. 2014.

[8] D. Streitfeld, "Fire phone immerses users in Amazon's world", The New York Times, June 2014, Online: http://www.nytimes.com/2014/06/19/technology/amazon-introduces-fire-smartphone.html.

[9] H. Li and M. Flierl, "Mobile 3D visual search using the Helmert transformation of stereo features", in *Proceedings of the IEEE International Conference on Image Processing*, Melbourne, Australia, Sept. 2013.

[10] X. Lyu, H. Li, and M. Flierl, "Hierarchically structured multi-view features for mobile visual search", in *Proceedings of the Data Compression Conference*, Snowbird, UT, Mar. 2014.

[11] D. Nister and H. Stewenius, "Scalable recognition with a vocabulary tree", in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, New York, NY, June 2006.

[12] D. Liu and M. Flierl, "Motion-adaptive transforms based on the Laplacian of vertex-weighted graphs", in *Proceedings of the Data Compression Conference*, Snowbird, UT, Mar. 2014.

[13] M. Fischler and R. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography", *Communications of the ACM*, vol. 24, no. 6, pp. 381–395, June 1981.

[14] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge University Press, 2nd edition, 2004.

[15] D. Ebri Mars, H. Wu, H. Li, and M. Flierl, "Joint geometric verification and ranking using multi-view vocabulary trees for mobile 3D visual search", in *Proceedings of the Data Compression Conference*, Snowbird, UT, Apr. 2015.

Interdisciplinarity

My application is interdisciplinary

☐

An interdisciplinary research project is defined in this call for proposals as a project that can not be completed without knowledge, methods, terminology, data and researchers from more than one of the Swedish Research Councils subject areas; Medicine and health, Natural and engineering sciences, Humanities and social sciences and Educational sciences. If your research project is interdisciplinary according to this definition, you indicate and explain this here.

Click here for more information

Scientific report

Scientific report/Account for scientific activities of previous project

3D Scenes in Adaptive Orthogonal Signal Spaces for Immersive Networked Experience
VR Project Research Grant
Registration number: 2011-5841
Period: 2012-1-1 until 2014-12-31
Amount: 2,293,000 SEK
One PhD student: Du Liu

Publications:

[6] Du Liu and Markus Flierl, Energy Compaction on Graphs for Motion-Adaptive Transforms, Proc. IEEE Data Compression Conference, Snowbird, UT, Apr. 2015, to appear.

[5] Du Liu and Markus Flierl, Motion-Adaptive Transforms based on the Laplacian of Vertex-Weighted Graphs, Proc. IEEE Data Compression Conference, Snowbird, UT, Mar. 2014.

[4] Du Liu and Markus Flierl, Graph-Based Construction and Assessment of Motion-Adaptive Transforms, Proc. Picture Coding Symposium, San Jose, CA, Dec. 2013.

[3] Du Liu and Markus Flierl, Graph-Based Rotation of the DCT Basis for Motion-Adaptive Transforms, Proc. IEEE International Conference on Image Processing, Melbourne, Australia, Sep. 2013.

[2] Du Liu and Markus Flierl, Motion-Adaptive Transforms based on Vertex-Weighted Graphs, Proc. IEEE Data Compression Conference, Snowbird, UT, Mar. 2013.

[1] Du Liu and Markus Flierl, Video Coding with Adaptive Motion-Compensated Orthogonal Transforms, Proc. Picture Coding Symposium, Krakow, Poland, May 2012.

Scientific results:

- Developed a novel algorithms for motion-adaptive transform coding of video and multi-view video that permit rate-distortion efficient and network friendly (i.e. embedded) data representations.

- Developed novel approaches to graph-based video processing

- Analyzed the relation between graph-based frequency analysis and graph-based energy compaction

- Developed new statistical models for graph-based video data to enhance the rate-distortion efficiency of graph-based video compression algorithms.

Relationship to current application: none

Budget and research resources

Project staff

Describe the staff that will be working in the project and the salary that is applied for in the project budget. Enter the full amount, not in thousands SEK.

Participating researchers that accept an invitation to participate in the application will be displayed automatically under Dedicated time for this project. Note that it will take a few minutes before the information is updated, and that it might be necessary for the project leader to close and reopen the form.

Dedicated time for this project

	Role in the project	Name	Percent of full time
1	Applicant	Markus Flierl	40
2	Other personnel without doctoral degree	PhD student	100

Salaries including social fees

	Role in the project	Name	Percent of salary	2016	2017	2018	2019	Total
1	Applicant	Markus Flierl	40	424,200	436,900	450,000	463,400	1,774,500
2	Other personnel without doctoral degree	PhD student	100	509,800	550,300	614,100	682,000	2,356,200
	Total			934,000	987,200	1,064,100	1,145,400	4,130,700

Other costs

Describe the other project costs for which you apply from the Swedish Research Council. Enter the full amount, not in thousands SEK.

Premises

	Type of premises	2016	2017	2018	2019	Total
1	Office space, EES/KT	113,000	119,000	129,000	139,000	500,000
	Total	113,000	119,000	129,000	139,000	500,000

Running Costs

	Running Cost	Description	2016	2017	2018	2019	Total
1	Conferences	Travel and fees	25,000	75,000	75,000	75,000	250,000
2	Equipment	Workstations, laptops	80,000		40,000		120,000
3	Publication costs	fees, posters	2,000	2,000	2,000	2,000	8,000
	Total		107,000	77,000	117,000	77,000	378,000

Depreciation costs

Depreciation cost	Description	2016	2017	2018	2019

Total project cost

Below you can see a summary of the costs in your budget, which are the costs that you apply for from the Swedish Research Council. Indirect costs are entered separately into the table.

Under Other costs you can enter which costs, aside from the ones you apply for from the Swedish Research Council, that the project includes. Add the full amounts, not in thousands of SEK.

The subtotal plus indirect costs are the total per year that you apply for.

Total budget

Specified costs	2016	2017	2018	2019	Total, applied	Other costs	Total cost
Salaries including social fees	934,000	987,200	1,064,100	1,145,400	4,130,700		4,130,700
Running costs	107,000	77,000	117,000	77,000	378,000		378,000
Depreciation costs					0		0
Premises	113,000	119,000	129,000	139,000	500,000		500,000
Subtotal	1,154,000	1,183,200	1,310,100	1,361,400	5,008,700	0	5,008,700
Indirect costs	342,000	361,000	389,000	418,000	1,510,000		1,510,000
Total project cost	1,496,000	1,544,200	1,699,100	1,779,400	6,518,700	0	6,518,700

Explanation of the proposed budget

Briefly justify each proposed cost in the stated budget.

Explanation of the proposed budget*

Salary for one phd student at 100% for four years: Direct costs 589k SEK per year.

Salary for PI at 40% for four years: Direct costs 444k SEK per year.

Travel costs to attend conferences for PhD student and PI: Total over four years: 250k SEK. (International conference: 25k SEK, European conference 12.5k SEK)

Equipment for four years: 120k SEK. Includes laptop and desk equipment for PhD student and PI; Workstations, server, and simulation computers to implement the project

Publication costs for four years: 8k SEK for page charges and posters

Office space at KTH, EES/KT Lab, for one PhD student (100%) and PI (40%) for four years: 500k SEK

Indirect costs for four years, calculated from overhead costs at KTH, EES/KT Lab: 1,510k SEK

Total project cost for four years: 6,519k SEK

Other funding

Describe your other project funding for the project period (applied for or granted) aside from that which you apply for from the Swedish Research Council. Write the whole sum, not thousands of SEK.

Other funding for this project

Funder	Applicant/project leader	Type of grant	Reg no or equiv.	2016	2017	2018	2019

CV and publications

CV

Markus Flierl, Associate Professor

Date of birth:	July 4th, 1971	Email:	mflierl@kth.se
Nationality:	German	Web:	https://people.kth.se/~mflierl/

1. Current positions at KTH

2009 – present	Associate professor, Communication Theory Laboratory, Visual Communication
2011 – present	Program director of the master program "Digital Media Technology" of the EIT ICT Labs Master School
2008 – present	Leader of the Visual Communication Group at the Communication Theory Laboratory

2. Education

1999 – 2003	Ph.D., Electrical Engineering, Friedrich Alexander University Erlangen-Nuremberg, Germany Thesis title: *Video Coding with Superimposed Motion-Compensated Signals* Research visit: Stanford University, 2000 – 2002 Main supervisor: Bernd Girod
1997	Diploma, Electrical Engineering, Friedrich Alexander University Erlangen-Nuremberg, Germany

3. Previous employment

2008 – 2009	Research Group Leader, School of Electrical Engineering, KTH, Sweden
2005 – 2008	Visiting Assistant Professor and Max Planck Research Group Leader, Department of Electrical Engineering, Stanford University, USA
2003 – 2005	Postdoctoral Researcher, Signal Processing Laboratories, EPFL École Polytechnique Fédérale de Lausanne, Switzerland
2000 – 2002	Research Assistant, Department of Electrical Engineering, Stanford University, USA
1999 – 2000	Research Assistant, Telecommunications Laboratory, Friedrich Alexander University Erlangen-Nuremberg, Germany
1997 – 1998	Community Service, University Hospital Erlangen, Germany

4. Selected awards and fellowships

2007	SPIE VCIP Young Investigator Award
2007	SPIE VCIP Best Paper Award
2006	IEEE MMSP Best Student Paper Award for supervised students
2005	Max Planck Fellowship and Research Grant
1999	Fellowship of the Graduate Research Center, Friedrich Alexander University, Erlangen-Nuremberg, Germany
1997	Electrical Engineering Diploma Degree Program Achievement Award, Friedrich Alexander University, Erlangen-Nuremberg, Germany

5. Supervision of students and postdoctoral fellows

2008 – present	2 postdocs, 4 PhD students, and 22 master students, School of Electrical Engineering, KTH, Sweden
2005 – 2008	2 PhD students, Department of Electrical Engineering, Stanford University, USA

6. Selected teaching

2010 – present	EN2401 Image and Video Processing, KTH
2011 – present	EN2500 Information Theory and Source Coding, KTH
2010 – present	EN2600 Project Course on Multimedia Signal Processing, KTH
2006 – 2008	EE398 Image and Video Compression (co-instructor), Stanford University

7. Selected commissions of trust

2014	Expert for the Swiss National Science Foundation
2012 – 2014	Member of the ACCESS executive committee at KTH
2011 – present	Program director of the master program “Digital Media Technology” of the EIT ICT Labs Master School
2008 – present	Member of doctoral examination committees
2006	Consultant for the Institute for the Promotion of Innovation by Science and Technology in Flanders, Belgium
2003 – present	Member of Technical Program Committees

8. Memberships of scientific societies

2000 – present	Member, IEEE Signal Processing Society

9. Major collaborations

2014 – present	Peter Eisert, Image-based Features, Heinrich Hertz Institute, Fraunhofer, Germany
2011 – present	Erik Jansen, Faculty of Electrical Engineering, Mathematics and Computer Science, Delft University of Technology, Netherlands
2011 – present	Tim Weyrich, Department of Computer Science, University College London, UK

10. Grants

2014	EIT ICT Labs, “Mobile 3D Visual Search,” 250k EUR budget
2012 – 2015	Swedish Research Council, “3D Scenes in Adaptive Orthogonal Signal Spaces for Immersive Networked Experience,” 1 Ph.D. student
2010 – 2013	EU FP7, “Free-Viewpoint Immersive Networked Experience,” 1 Ph.D. student
2008 – 2011	Ericsson Research, “3D Video,” 1 Ph.D. student
2005 – 2008	Max Planck Research Grant “Visual Sensor Networks,” 2 Ph.D. students

Publication list

1 Selected Papers in Journals

[1] R. Zhi, M. Flierl, Q. Ruan, and W.B. Kleijn, "Graph-preserving sparse nonnegative matrix factorization with application to facial expression recognition", *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 41, no. 1, pp. 38–52, Feb. 2011. [Cited by 66]

[2] M. Flierl and P. Vandergheynst, "Distributed coding of highly correlated image sequences with motion-compensated temporal wavelets", *EURASIP Journal on Applied Signal Processing, Special Issue on Video Analysis and Coding for Robust Transmission*, vol. 2006, ID 46747, 2006. [Cited by 12]

[3] Z. Ma, P.K. Rana, J. Taghia, M. Flierl, and A. Leijon, "Bayesian estimation of Dirichlet mixture model with variational inference", *Pattern Recognition*, vol. 47, no. 9, pp. 3143–3157, Sep. 2014. [Cited by 3]

[4] P.K. Rana, J. Taghia, Z. Ma, and M. Flierl, "Probabilistic multiview depth image enhancement using variational inference", *IEEE Journal of Selected Topics in Signal Processing*, vol. 9, no. 2, Apr. 2015. [Cited by 1]

Selected Papers in Conference Proceedings

[1] M. Flierl and P. Vandergheynst, "An improved pyramid for spatially scalable video coding", in *Proceedings of the IEEE International Conference on Image Processing*, Genova, Italy, Sep. 2005. [Cited by 30]

[2] M. Flierl, "Video coding with lifted wavelet transforms and frame-adaptive motion compensation", in *Visual Content Processing and Representation*, N. García, J.M. Martínez, and L. Salgado, Eds., vol. 2849 of *Lecture Notes in Computer Science*, pp. 243–251. Springer-Verlag, Berlin, 2003, Proceedings of the 8th International Workshop VLBV, Madrid, Spain, Sep. 2003. [Cited by 25]

[3] P.K. Rana and M. Flierl, "Depth consistency testing for improved view interpolation", in *Proceedings of the IEEE International Workshop on Multimedia Signal Processing*, Saint-Malo, France, Oct. 2010. [Cited by 11]

[4] M. Flierl, "Adaptive spatial wavelets for motion-compensated orthogonal video transforms", in *Proceedings of the IEEE International Conference on Image Processing*, Cairo, Egypt, Nov. 2009. [Cited by 10]

[5] H. Li and M. Flierl, "SIFT-based multi-view cooperative tracking for soccer video", in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, Kyoto, Japan, Mar. 2012. [Cited by 6]

[6] H. Li and M. Flierl, "Rate-distortion-optimized content-adaptive coding for immersive networked experience of sports events", in *Proceedings of the IEEE International Conference on Image Processing*, Brussels, Belgium, Sep. 2011. [Cited by 6]

[7] R. Zhi, M. Flierl, Q. Ruan, and W.B. Kleijn, "Facial expression recognition based on graph-preserving sparse non-negative matrix factorization", in *Proceedings of the IEEE International Conference on Image Processing*, Cairo, Egypt, Nov. 2009. [Cited by 5]

[8] M. Flierl and P. Vandergheynst, "Video coding with motion-compensated temporal transforms and side information", in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, Philadelphia, PA, Mar. 2005. [Cited by 5]

[9] M. Flierl and P. Vandergheynst, "Inter-resolution transform for spatially scalable video coding", in *Proceedings of the Picture Coding Symposium*, San Francisco, CA, Dec. 2004. [Cited by 5]

[10] H. Helgason, H. Li, and M. Flierl, "Multiscale framework for adaptive and robust enhancement of depth in multi-view imagery", in *Proceedings of the IEEE International Conference on Image Processing*, Orlando, FL, Sep. 2012. [Cited by 4]

For a complete list of more than 77 publications, please visit
http://people.kth.se/~mflierl/publications.html

Please see also Google Scholar
http://scholar.google.com/citations?user=yg9ZalsAAAAJ&hl=en

2 Research Monographs

[1] M. Flierl and B. Girod, *Video Coding with Superimposed Motion-Compensated Signals: Applications to H.264 and Beyond*, Kluwer Academic Publishers / Springer, Boston, 2004.

3 Granted Patents

[1] M. Flierl and B. Girod, "Systems, methods, devices and arrangements for motion-compensated image processing and coding ", US patent 8,346,000, Jan. 2013.

[2] M. Flierl, B. Girod, and P. Vandergheynst, "Image transform for video coding", US patent 8,300,693, Oct. 2012.

[3] M. Flierl and P. Vandergheynst, "Method for spatially scalable video coding", US patent 7,616,824, Nov. 2009.

[4] T. Wiegand, M. Flierl, and B. Girod, "Multi-hypothesis motion-compensated video image predictor", US patent 6,807,231, Oct. 2004.

4 Selected Invited Presentations

[1] X. Lyu, H. Li, and M. Flierl, "Hierarchically Structured Multi-view Features for Mobile Visual Search", *IEEE Data Compression Conference*, Snowbird, UT, Mar. 2014.

[2] M. Flierl, "Motion-compensated Orthogonal Transforms for Multiview Video Coding", *European Signal Processing Conference - EUSIPCO*, Poznan, Poland, Sep. 2007.

[3] M. Flierl, "Coding Efficiency of Video Sensor Networks", *IEEE International Conference on Image Processing*, Genova, Italy, Sep. 2005.

[4] M. Flierl, "From Motion Pictures to Moving 3D Scenes: New Concepts for Image-Based Communication", University of California Los Angeles, Los Angeles, CA, Feb. 5, 2008.

[5] M. Flierl, "Motion-Compensated Orthogonal Video Transforms", University of Southern California, Los Angeles, CA, Feb. 4, 2008.

[6] M. Flierl, "Video Coding with Motion-Compensated Lifted Wavelet Transforms", University of Washington, Seattle, WA, Dec. 13, 2004.

5 International Awards

SPIE VCIP Young Investigator Award, 2007.
SPIE VCIP Best Paper Award, 2007.
IEEE MMSP Best Student Paper Award for supervised students, 2006.
Max Planck Fellowship and Research Grant, 2005.

CV

Name: Markus Flierl	**Doctorial degree:** 2003-06-27
Birthdate: 19710704	**Academic title:** Docent
Gender: Male	**Employer:** Kungliga Tekniska högskolan

Research education

Dissertation title (swe)
Video Coding with Superimposed Motion-Compensated Signals

Dissertation title (en)
Video Coding with Superimposed Motion-Compensated Signals

Organisation	Unit	Supervisor
Friedrich Alexander University Erlangen-Nuremberg, Germany Not Sweden - Higher Education institutes	Telecommunications Laboratory	Bernd Girod

Subject doctors degree	ISSN/ISBN-number	Date doctoral exam
20204. Telekommunikation	978-1402077593	2003-06-27

Publications

Name: Markus Flierl	**Doctorial degree:** 2003-06-27
Birthdate: 19710704	**Academic title:** Docent
Gender: Male	**Employer:** Kungliga Tekniska högskolan

Flierl, Markus has not added any publications to the application.

Register

Terms and conditions

The application must be signed by the applicant as well as the authorised representative of the administrating organisation. The representative is normally the department head of the institution where the research is to be conducted, but may in some instances be e.g. the vice-chancellor. This is specified in the call for proposals.

The signature*from the applicant*confirms that:

- the information in the application is correct and according to the instructions form the Swedish Research Council
- any additional professional activities or commercial ties have been reported to the administrating organisation, and that no conflicts have arisen that would conflict with good research practice
- that the necessary permits and approvals are in place at the start of the project e.g. regarding ethical review.

The signature *from the administrating organisation* confirms that:

- the research, employment and equipment indicated will be accommodated in the institution during the time, and to the extent, described in the application
- the institution approves the cost-estimate in the application
- the research is conducted according to Swedish legislation.

The above-mentioned points must have been discussed between the parties before the representative of the administrating organisation approves and signs the application.

Project out lines are not signed by the administrating organisation. The administrating organisation only sign the application if the project outline is accepted for step two.

Applications with an organisation as applicant is automatically signed when the application is registered.