ILMENAU UNIVERSITY OF TECHNOLOGY

fayteq
technologies beyond imagination

Qualcomm Augmented Reality Lecture Series

The Magic of Diminished Reality

Real-time Video Inpainting

Jan Herling

CV

- **2002-2008:**
  Study of computer science at RWTH Aachen University, Germany
- **2002-2008:**
  Student assistant at the Fraunhofer Institute for Applied Information Technology FIT, Sankt Augustin, Germany
- **2008-2009:**
  Research assistant at FIT, Participation in European Projects: IPCity, Cospaces, EXPLOAR
- **since 2009:**
  Research assistant and PhD student at Ilmenau University of Technology, Germany
- **since 2012:**
  CTO and co-founder of fayteq GmbH

Outline

- Introduction
- Previous Research
- Real-time Video Inpainting
- What’s coming next?
- Live Demo

Introduction

Augmented Reality (AR)

- Combining reality and virtual content
- Augmenting the reality with virtual content
- Real-time application

Introduction

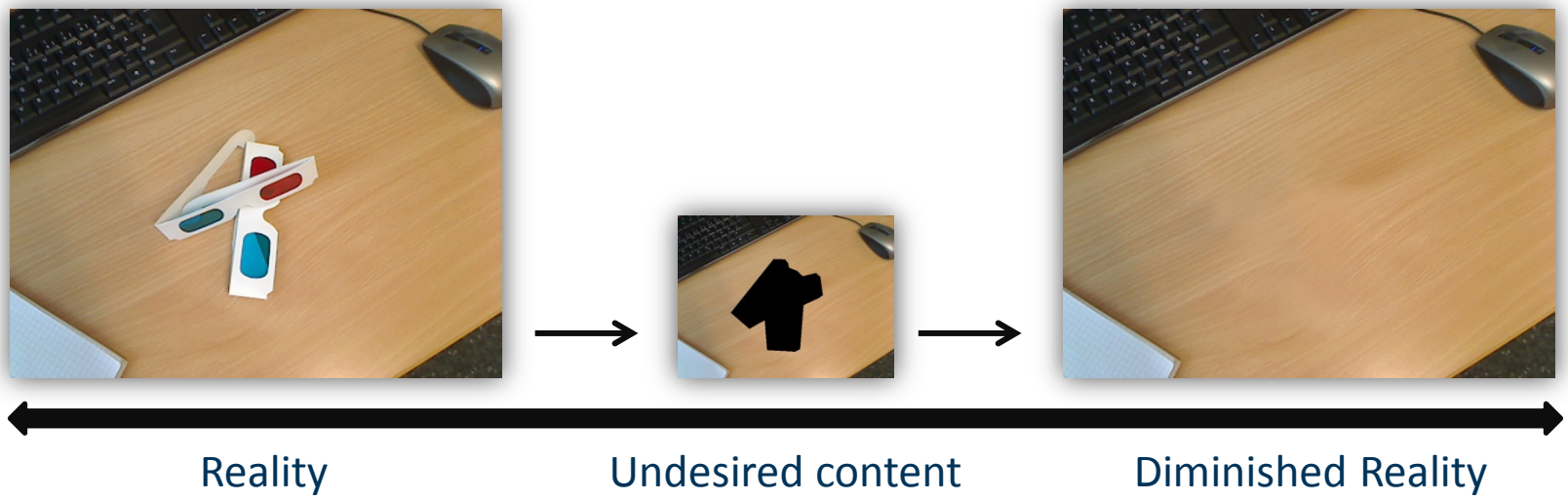
Augmented Reality (AR)

- Applications

Introduction

Diminished Reality (DR)

- Removing real objects from the reality
- Real-time application



Introduction

Diminished Reality (DR)

- Applications

Introduction

Diminished Reality (DR)

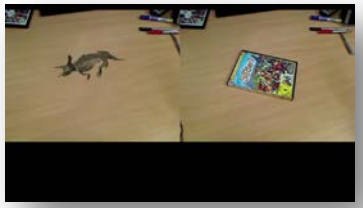
- Applications

Original images: Dr. Ulrich Heide, kindly authorized by the photographer

Introduction

Mediated Reality

- Combination of Augmented (or Mixed) and Diminished Reality
- Definition by Steven Mann in 1994:
  - A reality without constraints
  - Augmenting, enhancing, diminishing or altering the reality in an arbitrary way
  - Real-time application



Introduction

Mediated Reality

- Applications for architects

Images: Michael Lippert, kindly authorized by the author

Introduction

Mediated Reality

- Gaming applications

Images: Michael Lippert, kindly authorized by the author

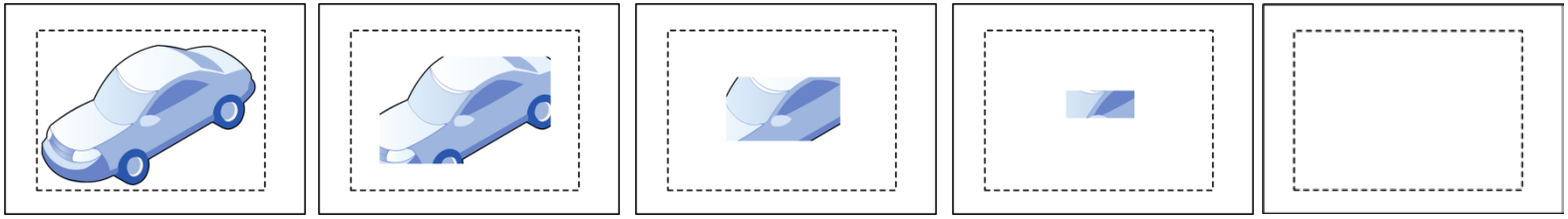
Outline

- Introduction
- Previous Research
  - Image Inpainting
  - Diminished Reality
- Real-time Video Inpainting
- What's coming next?
- Live Demo

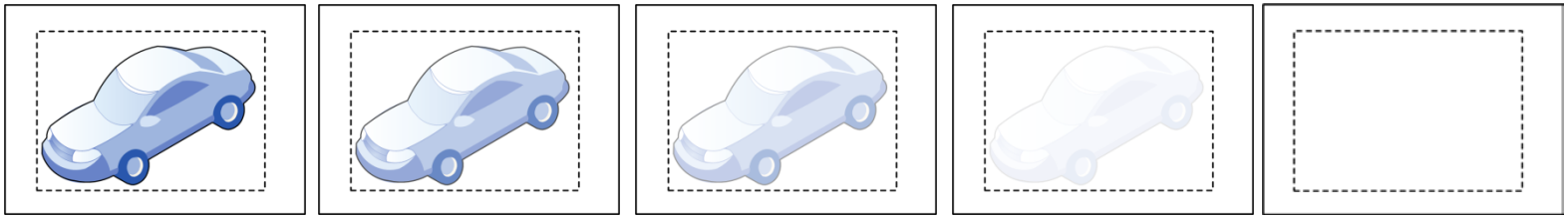
Previous Research

Image Inpainting / Image Completion

- Shrinking approaches



- Vanishing approaches



Previous Research

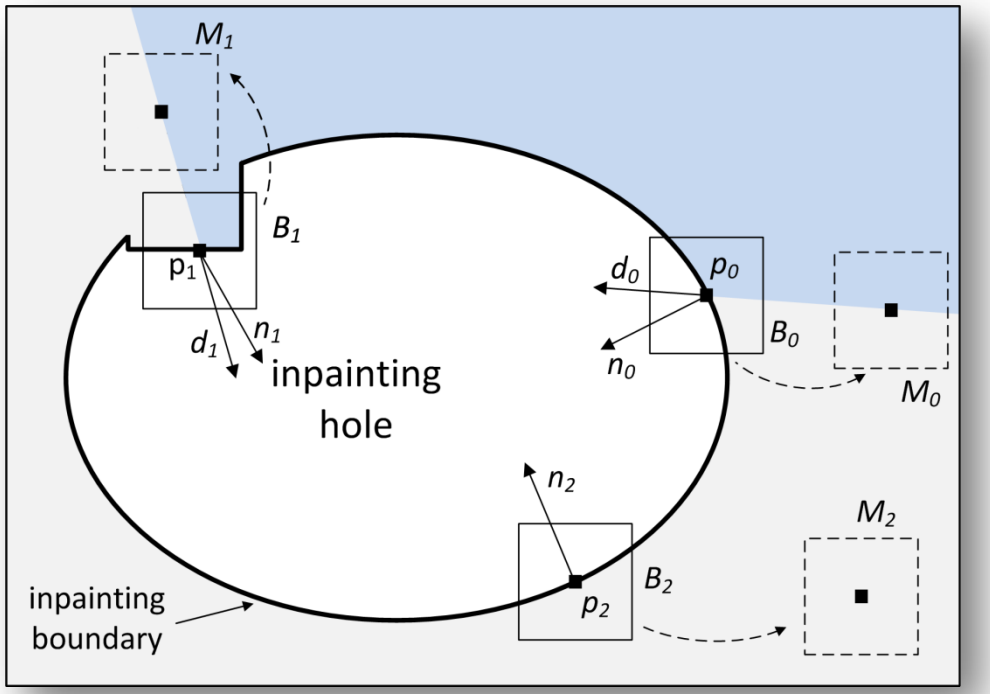
Image Inpainting

- Drori et al., 2003
- Separation of source and target region
- Patch-based inpainting
- Diffusion of inpainting boundary
- Inpainting by scale and orientation invariant image patches

Previous Research

Image Inpainting

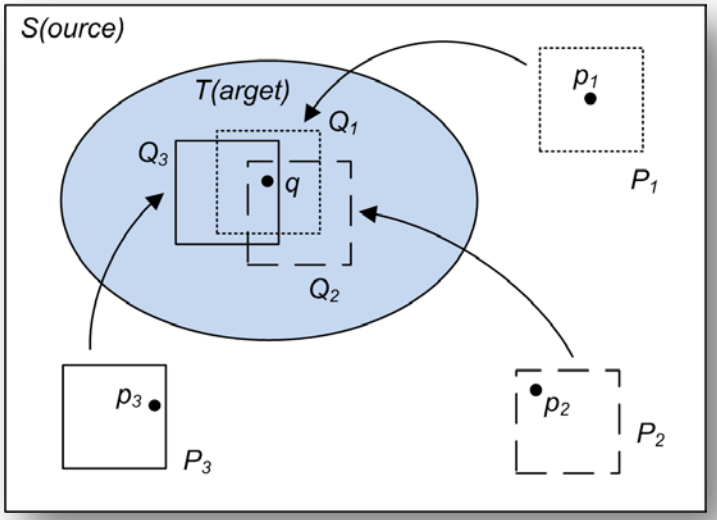
- Criminisi et al., 2003
- Patch-based inpainting
- Shrinking of the inpainting mask
- Inpainting priority mechanism



Previous Research

Image Inpainting

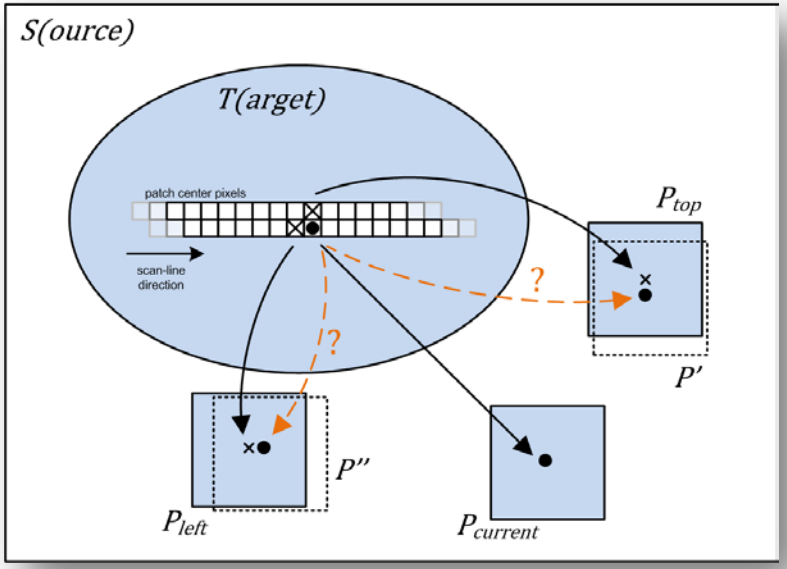
- Wexler et al., 2007
- Patch-based inpainting
- Iterative optimization approach



Previous Research

Image Inpainting

- Barnes et al., 2009
- First “interactive” image manipulation approach
- Patch-based inpainting
- Randomized optimization
- Mapping propagation



Previous Research

Diminished Reality

- Multi-view approaches
  - Application of several video cameras
  - Hiding of volumetric objects
  - Approx. 5 approaches
- Single-view approaches
  - One handheld camera
  - Synthesis of unknown image content
  - Often for marker hiding
  - Approx. 4 approaches

Previous Research

Diminished Reality, Multi-view Approaches

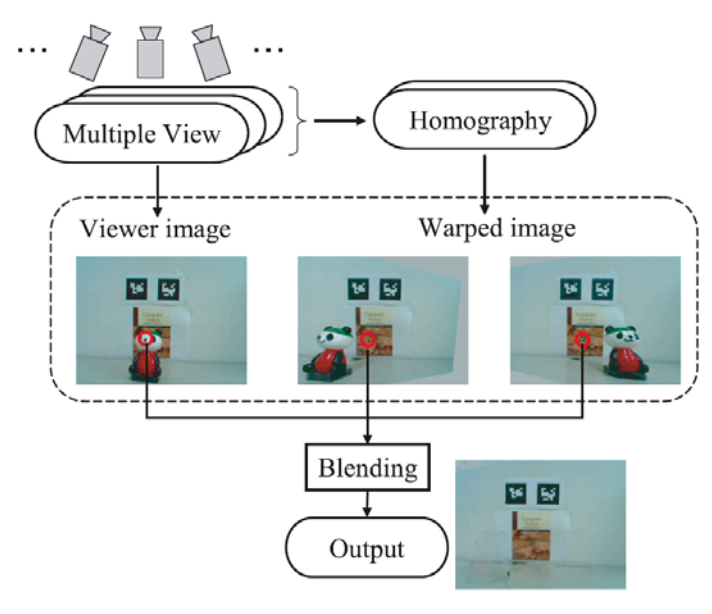
- Zokai et al., 2003
- Several registered reference images are applied to remove an undesired object

Multiview Paraperspective Projection Model for Diminished Reality, Zokai, Siavash and Esteve, Julien and Genc, Yakup and Navab, Nassir Proceedings of the 2nd IEEE/ACM International Symposium on Mixed and Augmented Reality, 2003

Previous Research

Diminished Reality, Multi-view Approaches

- Enomoto and Saito, 2007
- Several hand-held live cameras are applied to remove an undesired object



Diminished Reality using Multiple Handheld Cameras,
Akihito Enomoto and Hideo Saito, ACCV'07 Workshop on
Multi-dimensional and Multi-view Image Processing, Tokyo, Nov., 2007

Previous Research

Diminished Reality, Single-view approaches

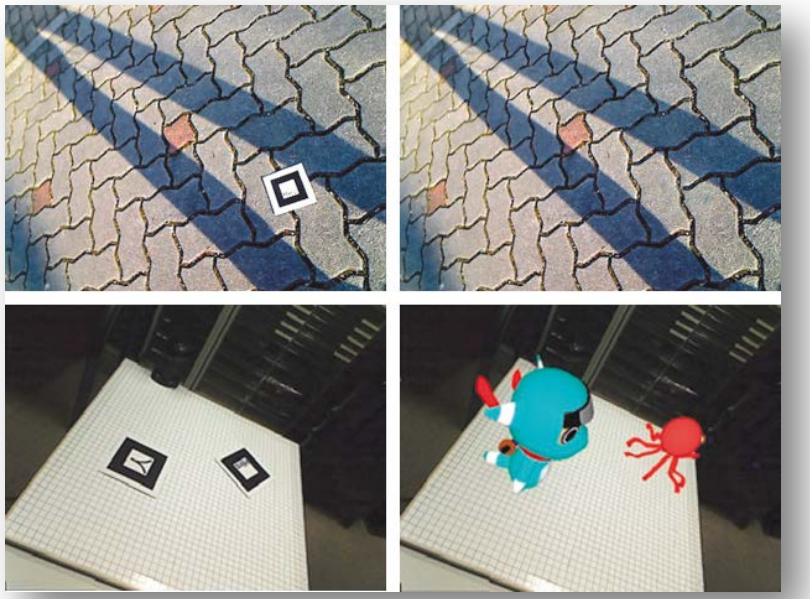
- Siltanen, 2006
- Marker hiding with synthesized texture
- Image synthesis by mirroring visual information

Texture generation over the marker area, Sanni Siltanen, Proceedings of the 5th IEEE and ACM International Symposium on Mixed and Augmented Reality, ISMAR 2006

Previous Research

Diminished Reality, Single-view approaches

- Kawai et al., 2012
- Marker hiding with synthesized (and rectified) texture
- Supporting cast shadows
- Not real-time capable image synthesis



AR Marker Hiding Based on Image Inpainting and Reflection of Illumination Changes, Norihiko Kawai, Masayoshi Yamasaki, Tomokazu Sato, Naokazu Yokoya, 2012 IEEE International Symposium on Mixed and Augmented Reality (ISMAR), 2012

Previous Research

Diminished Reality, Discussion

Single-view approaches	Multi-view approaches

Outline

- Introduction
- Previous Research
- Real-time Video Inpainting
  - Lessens Learned and Objective
  - Static Image Inpainting
  - Object Selection & Tracking
  - Video Inpainting
- What's coming next?
- Live Demo

Real-time Video Inpainting

Lessens Learned

- Limitations of the previous approach
  - Usage of grayscale image information
  - Sometimes blurred image results due to real-time constraints
- Requirements for an improved approach
  - Higher image qualities (but still real-time capable)
  - Avoiding of blurring image content
  - Explicit creation of a coherent video stream

Real-time Video Inpainting

Objective

- Single hand-held camera
- No a-priori information about the environment
- Selection of arbitrary objects
- Creation of a coherence video stream
- Real-time performance for Diminished Reality

Outline

- Introduction
- Previous Research
- Real-time Video Inpainting
  - Lessens Learned and Objective
  - Static Image Inpainting
  - Object Selection & Tracking
  - Video Inpainting
- What's coming next?
- Live Demo

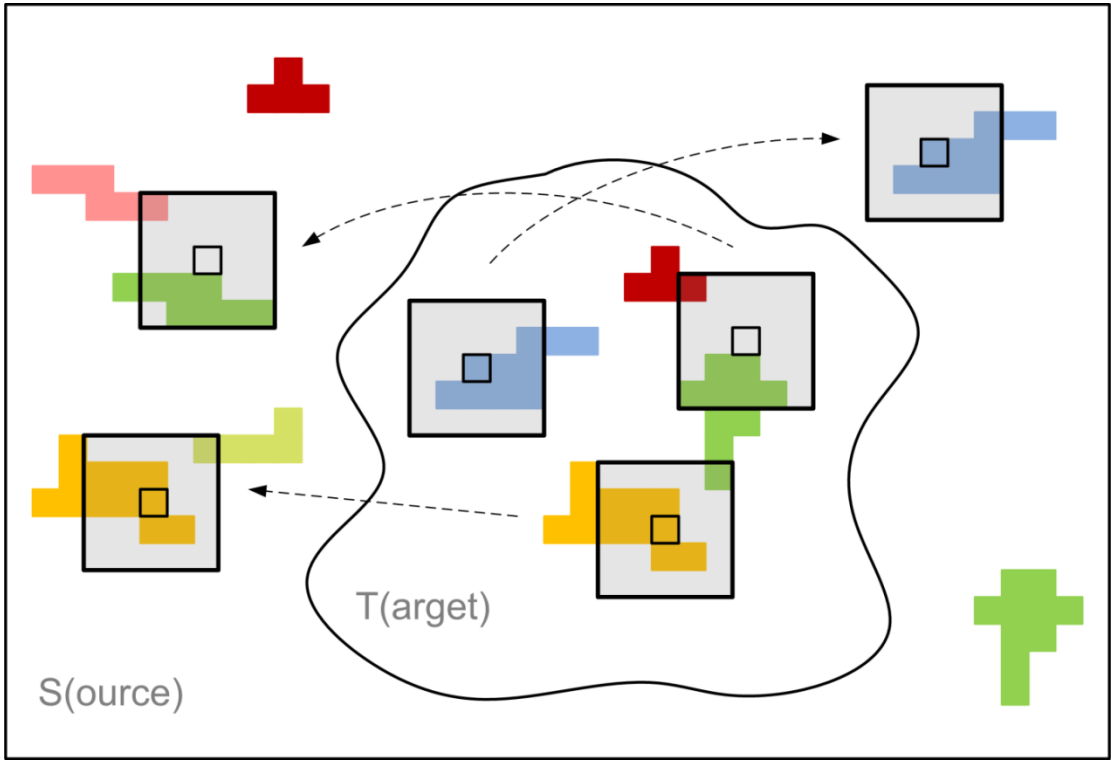
Static Image Inpainting

Approach Overview

- Separating image into **T**(arget) and **S**(ource) area
- Pixel-based image inpainting
  - Mapping f between **T**(arget) and **S**(ource) pixels
  - Copying pixels from remaining image content
- Combination of two cost constraints
  - Appearance mapping cost
  - Spatial mapping cost
- Cost minimization
  - Heuristic optimization approach
  - Multi-resolution optimization

Static Image Inpainting

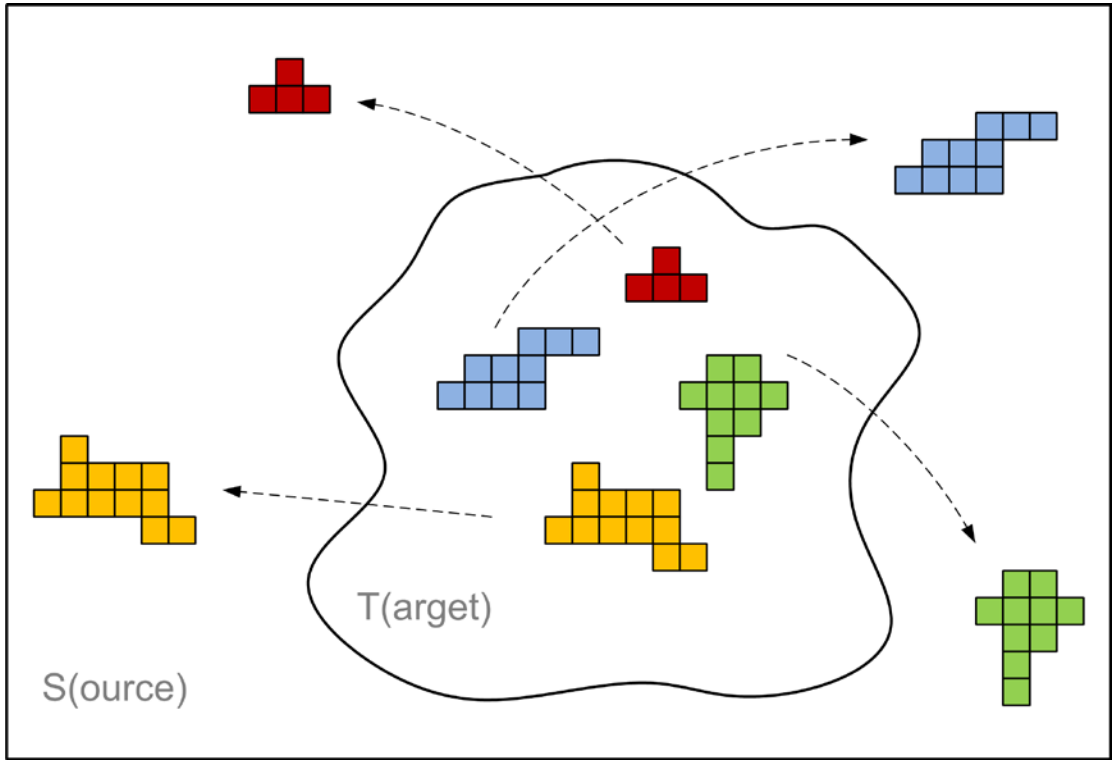
Appearance Mapping Cost



$$cost_{appearance}(p) = \sum_{\vec{v} \in N_a} d_a\left[i(p+\vec{v}), i(f(p)+\vec{v})\right] \cdot w_a(p+\vec{v})$$

Static Image Inpainting

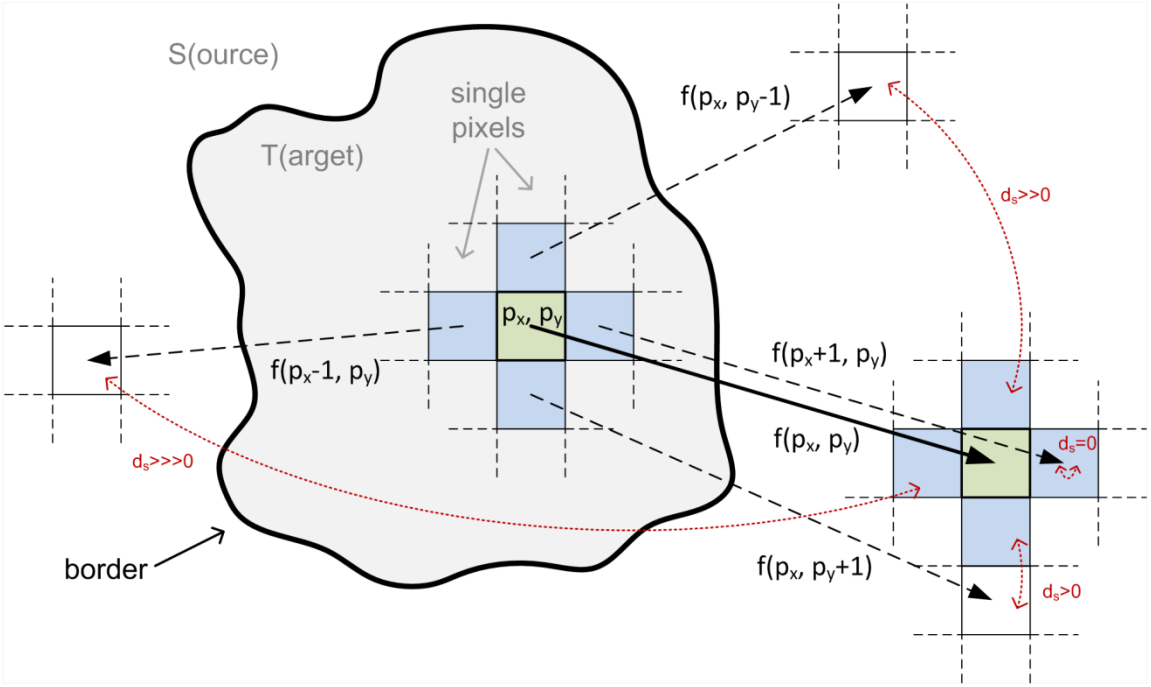
Spatial Mapping Cost



$$cost_{spatial}(p) = \sum_{\vec{v} \in N_s} d_s[f(p)+\vec{v}, f(p+\vec{v})] \cdot w_s(\vec{v})$$

Static Image Inpainting

Spatial Cost Constraint



$$cost_{spatial}(p) = \sum_{\vec{v} \in N_s} d_s[f(p)+\vec{v}, f(p+\vec{v})] \cdot w_s(\vec{v})$$

Static Image Inpainting

Pixel-Based Mapping Function

- Combination of both cost constraints

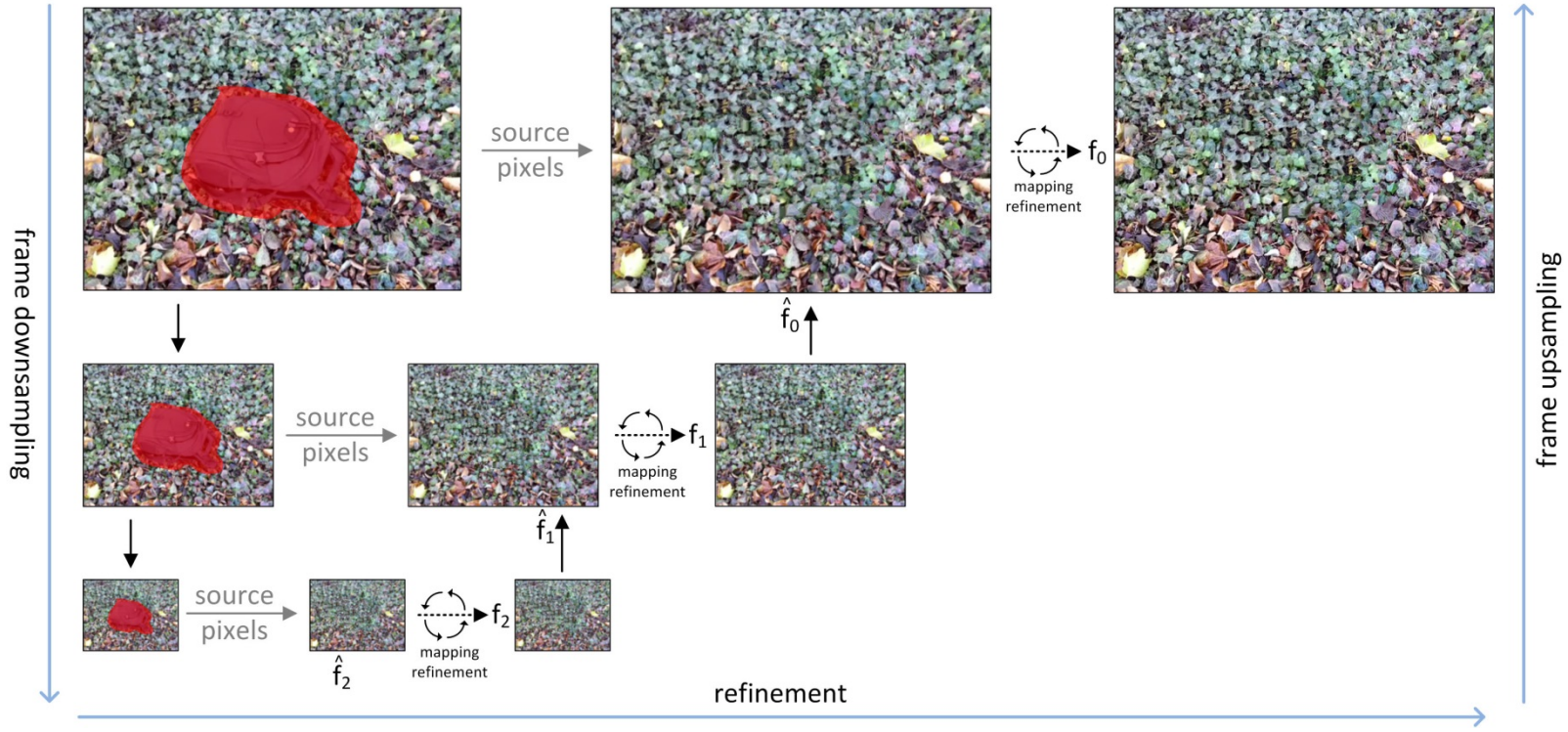
$$cost_{\alpha}(p) = \alpha \cdot cost_{spatial}(p) + (1-\alpha) \cdot cost_{appearance}(p)$$

- Finding mapping function f with minimal cost

$$\min_{f} \sum_{p \in T} cost_{\alpha}(p)$$

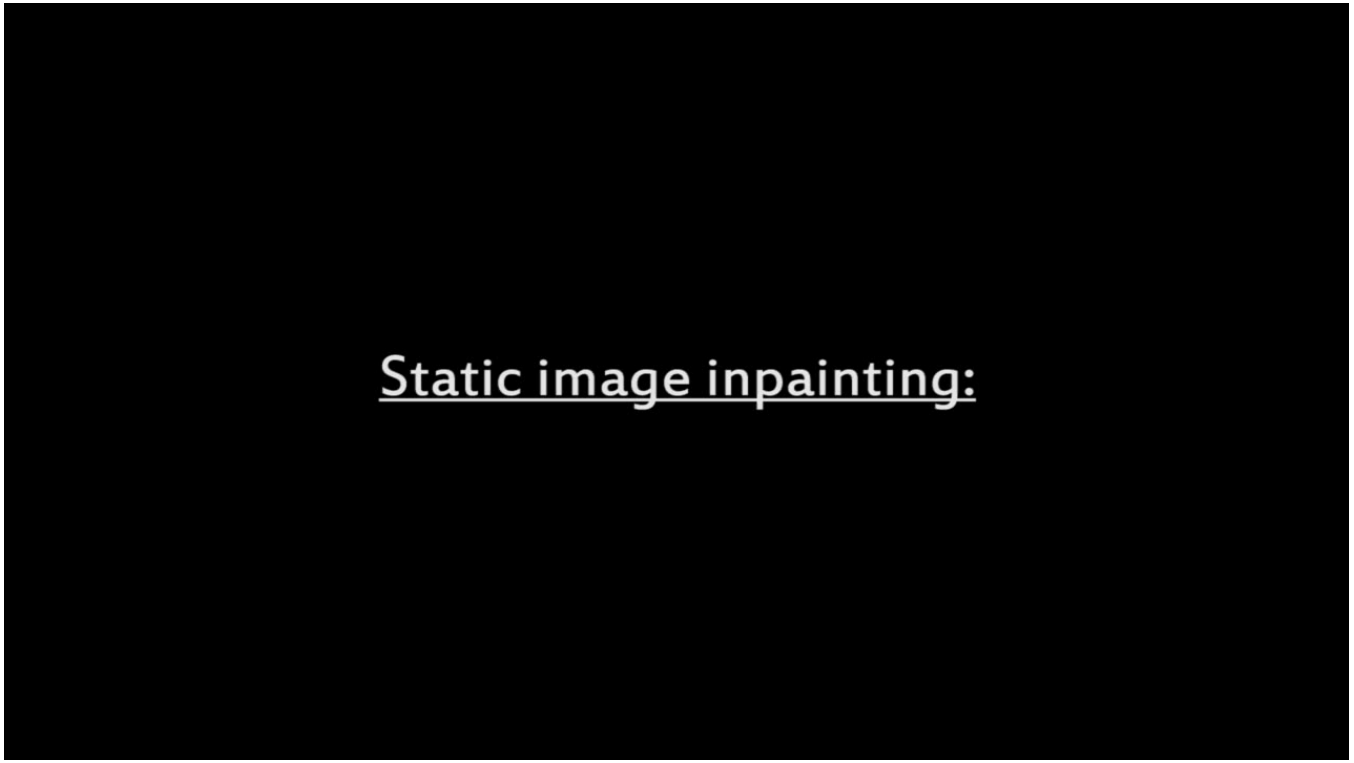
Static Image Inpainting

Multi-resolution Optimization



Static Image Inpainting

Individual Spatial Weighting



Static Image Inpainting

Results

Benchmark image (384x256),
13,048 inpainting pixels,
Drori et al., 2003

Xu et al, 2010,
> 4 minutes

Our result,
16.5 ms

Static Image Inpainting

Results

Benchmark image (538x403), 58,573 inpainting pixels, Criminisi et al., 2004

Kwok et al, 2010,
11,500.0 ms

Our result,
27.0 ms

Static Image Inpainting

Results

Original image: bbroianigo / pixelio.de,
(creative commons image database)

Image resolution: 1842x1266
182,860 inpainting pixels
Performance: **< 200 ms**

Static Image Inpainting

Properties

- Faster than our previous approach
  - Pixel-based inpainting avoids patch blending
  - Spatial cost allows faster convergence
- Resulting image quality is state-of-the-art
  - Supports textured image content
  - and supports structured image content
  - At least 100x faster than state-of-the-art approaches

Outline

- Introduction
- Previous Research
- Real-time Video Inpainting
  - Lessens Learned and Objective
  - Static Image Inpainting
  - Object Selection & Tracking
  - Video Inpainting
- What's coming next?
- Live Demo

Object Selection & Tracking

Requirements

- Application without knowledge about object or environment
- Real-time capability
- Simple user interaction
- Support for heterogeneous backgrounds

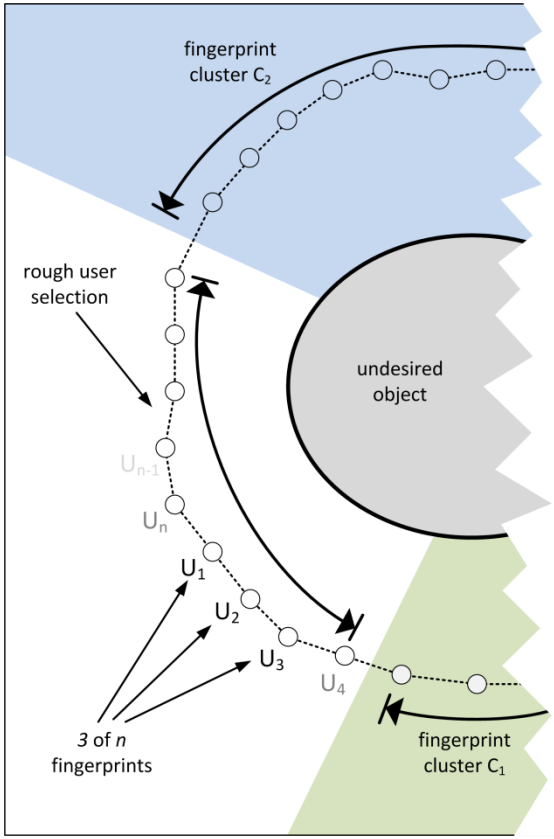
Object Selection & Tracking

Approach

- Selection
  - Fingerprints store visual characteristics
  - Comparing fingerprints with image content
  - Determination of a binary mask
- Tracking
  - Homography due to motion of contour points
  - New contour from homography
  - Contour refinement with fingerprints

Object Selection & Tracking

Rough User Selection

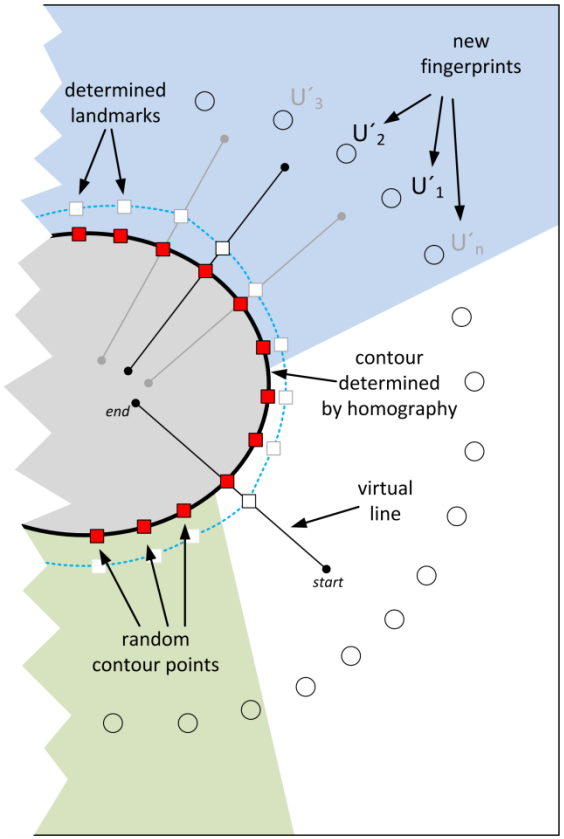


- Undesired object may have heterogeneous image background
- Rough user selection
- Distribution of fingerprints
- Clustering of fingerprints

- Maximal deviation in clusters provide reference threshold
- Comparison between fingerprints and image content provides undesired object
- Pixel is undesired if more than 95% of all fingerprints reject the pixel

Object Selection & Tracking

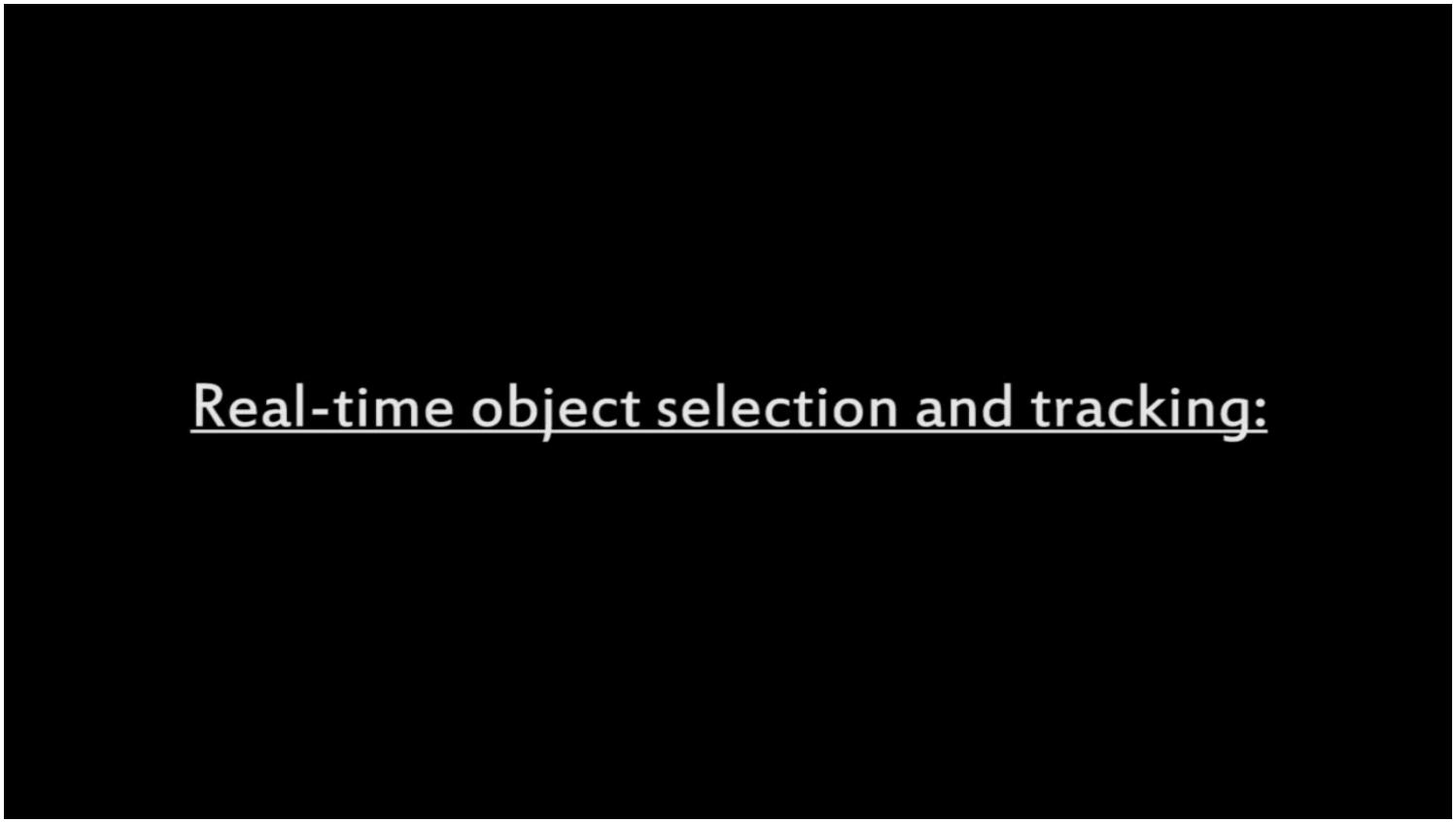
Homography Determination



- Rough contour from previous frame
- Homography determination at the object's contour
- New fingerprints (equally distributed)
- Subset of random contour points
- Virtual lines perpendicular to contour
- Determination of new landmarks
- Adjustment of the final contour

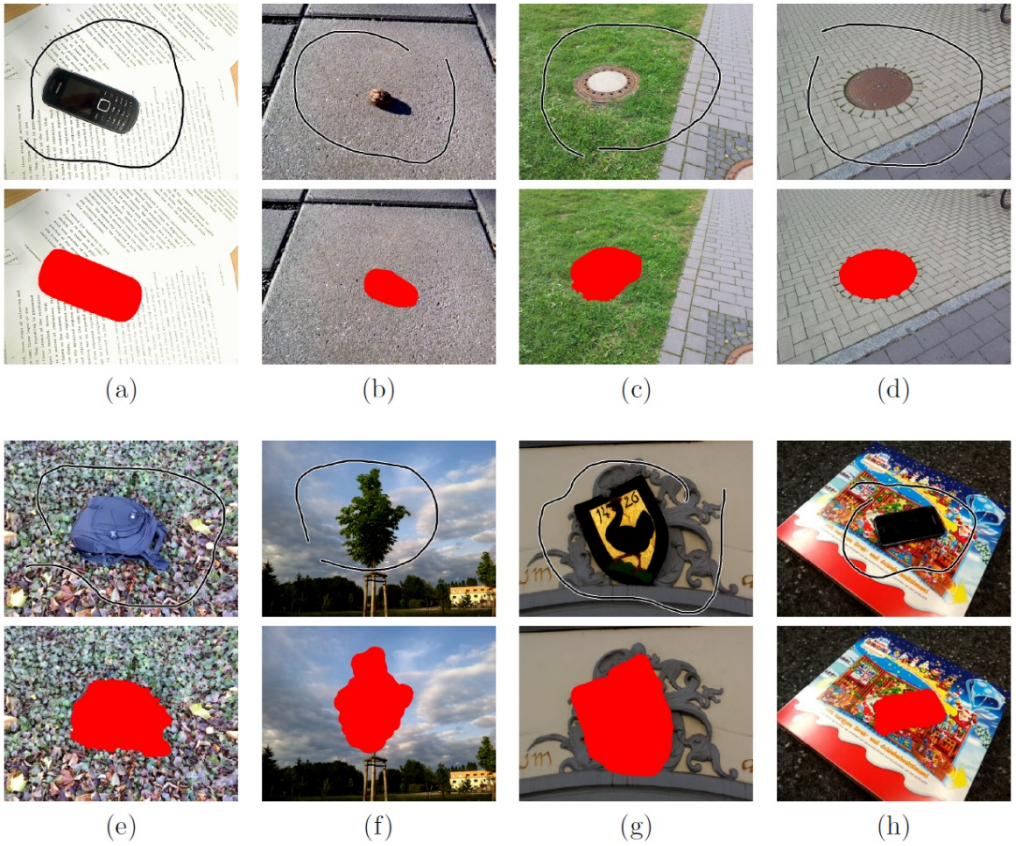
Object Selection & Tracking

Results



Object Selection & Tracking

Results



Outline

- Introduction
- Previous Research
- Real-time Video Inpainting
  - Lessens Learned and Objective
  - Static Image Inpainting
  - Object Selection & Tracking
  - Video Inpainting
- What's coming next?
- Live Demo

Video Inpainting

Video Coherence

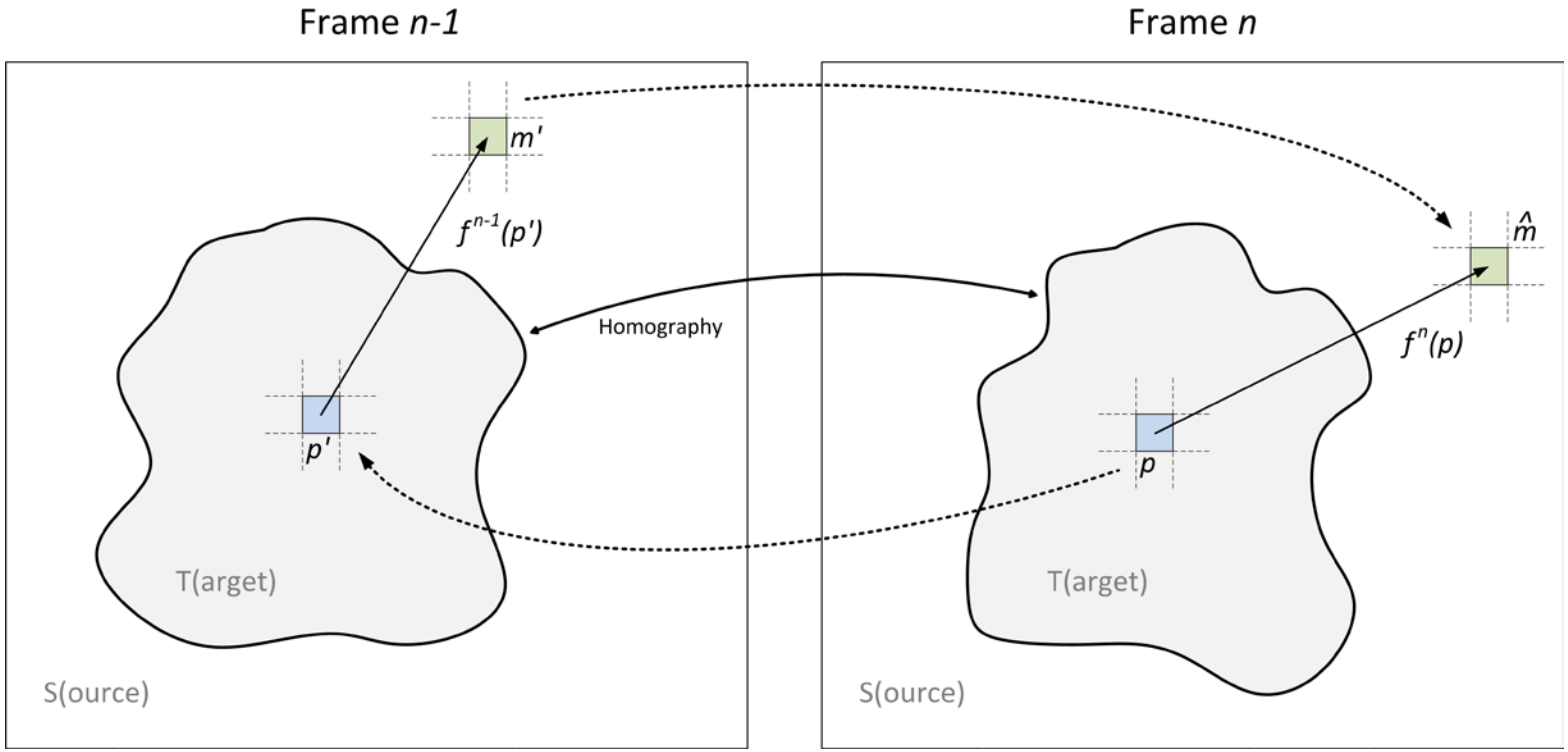
Video Inpainting

Requirements

- High image quality:
  - Image inpainting for each video frame
- High performance:
  - Information propagation from previous mapping
- Strong video coherence:
  - Application of a reference model

Video Inpainting

Information Propagation



Video Inpainting

Reference Model

- Warping of distinct key frames provide a visual model
- The model guides the inpainting process
- Extension of the appearance cost
  - Measure between model and inpainting result

$$cost'_{appearance}(p) = \sum_{\vec{v} \in N_a} d'_a[i(p+\vec{v}), r(p+\vec{v})] \cdot w'_a(p+\vec{v})$$

Video Inpainting

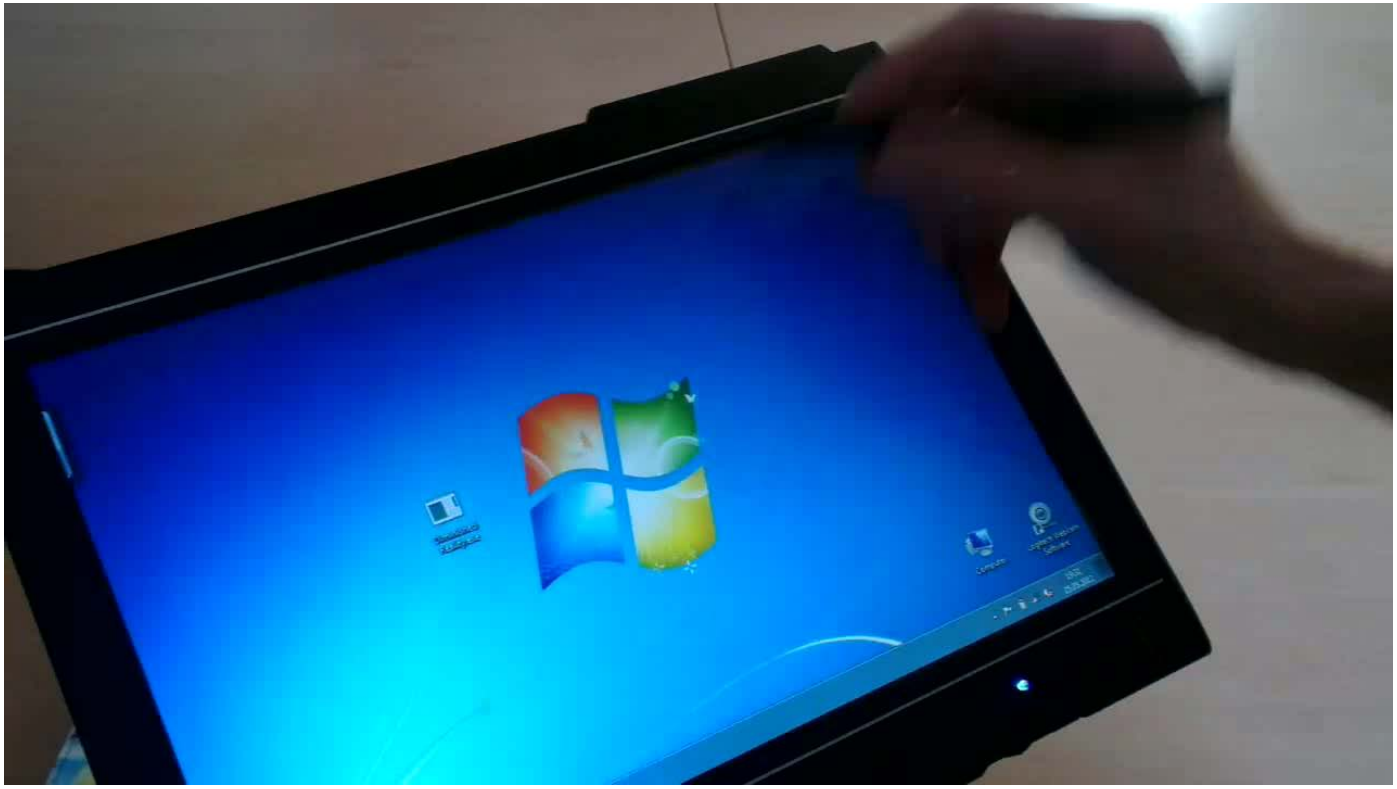
Performance

- Video resolution: 640x480 pixels
- Hardware: i7 mobile 2.13 GHz, 2010
- C++ Implementation running on CPU

	Object tracking	Reference model	Inpainting	Total
Removal mask: ~13K pixels, ~4.2%	4.62 ms	4.60 ms	11.16 ms	~ 50 fps
Removal mask: ~60K pixels, ~19.5%	5.71 ms	6.45 ms	53.37 ms	~ 15 fps

Video Inpainting

Results



Real-time Video Inpainting

Limitations

- The approach is suitable for
  - almost planar backgrounds
  - static objects
  - non-volumetric objects
- The approach fails to
  - remove volumetric objects with unconstraint camera motion
  - dynamic objects

Outline

- Introduction
- Previous Research
- Real-time Video Inpainting
- What’s coming next?
- Live Demo

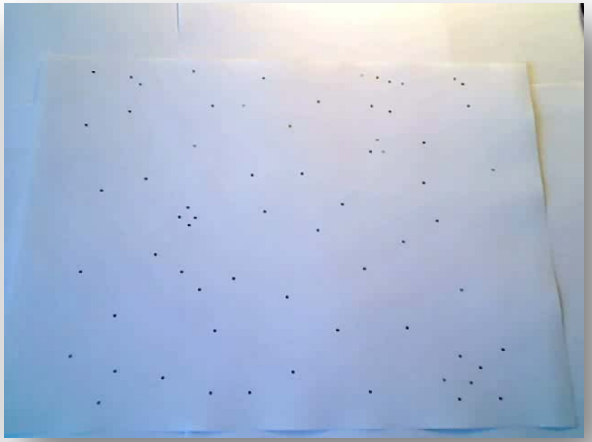
What's coming next?

- Support of mobile devices
  - The improved algorithm seems to be fast enough for the current smartphone generation
- Extension to non-planar backgrounds
  - High efficient tracking approaches
  - SLAM-like map creation of the environment
- Support for volumetric objects
  - Prediction of camera movements
- Investigation in GPU implementation

What's coming next?

High Efficient Tracking Approach

- The current inpainting pipeline
  - approx. 10% for tracking
  - approx. 90% for synthesis



What's coming next?

High Efficient Tracking Approach

- Tracking performance with random model variation

Platform	2010 i7 Laptop 2.13 GHz, 640x480	Samsung Galaxy S2, 320x240
Gray scale image	0.32ms	-
Feature detection	1.59ms (600 FAST features)	7.44ms (400 FAST features)
Pose determination	1.47ms	10.16ms
Overall	3.37ms	17.60ms

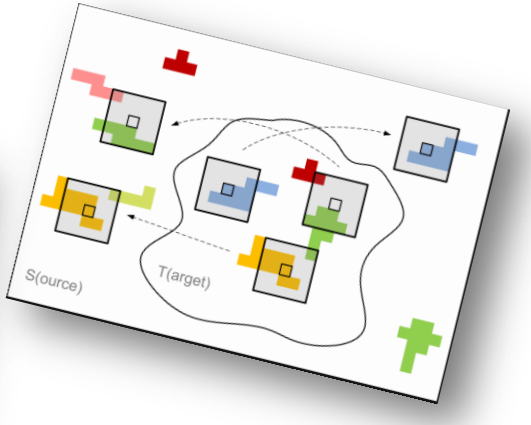
What's coming next?

Visions

- Powerful Mediated Reality applications
  - Combination of AR and Diminished Reality
  - Novel Mediated Reality Games
  - …

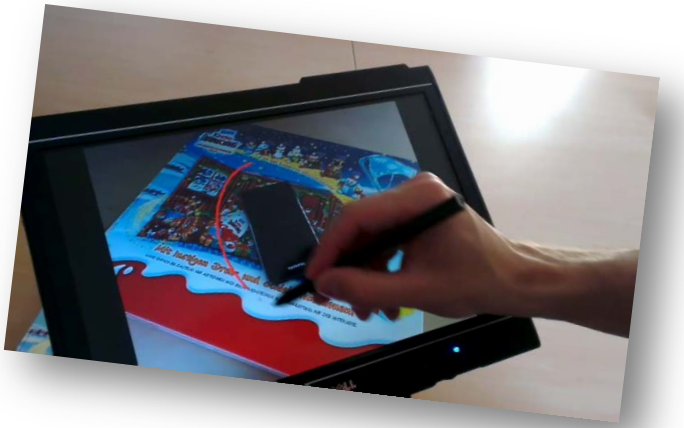
Outline

- Introduction
- Previous Research
- Real-time Video Inpainting
- What’s coming next?
- Live Demo



Thank you for your attention!

Further questions?



Contact:
jan.herling@fayteq.com
jan.herling@tu-ilmenau.de

fayteq
technologies beyond imagination