



Moritz Schmidt

SIFT Flow

Dense Correspondence across Different Scenes

Motivation

Bildregistrierung

- Bilder ähnlicher Szenen in Übereinstimmung bringen
- Transformation eines Bildes zu Referenzbild
- Optimierungsproblem

Wozu?

- Bilder nahtlos aneinandersetzen (Panoramen)
- Aufnahmen vergleichbar machen



Motivation

Probleme

- Verzerrungen
 - Beleuchtung
 - Hintergrund
- Objektunterschiede
 - Objekthäufigkeit



(i) Different perspectives and occlusions



(ii) Multiple objects; no global transform



(iii) Background clutter



(iv) High intra-class variations

Motivation

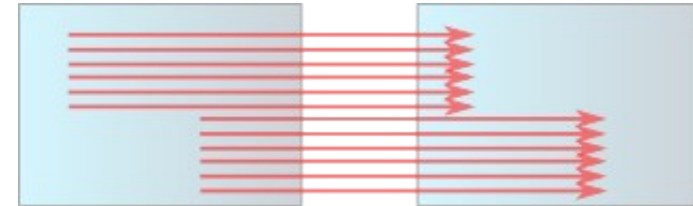
Ziel

- Ausrichten von Bildern trotz unterschiedlicher Szenen
- Robust gegenüber
 - Änderung der Perspektive, Beleuchtung
 - Fehlen oder Mehrfachvorkommen eines Objekts
- Endprodukt: Verschiebungsvektor für jeden Pixel

SIFT Flow vs Optical Flow

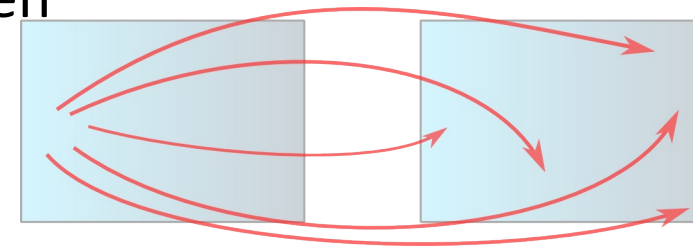
Optischer Fluss

- Beleuchtung ähnlich, Unterschiede gering
- Dichtes Verschiebungsfeld



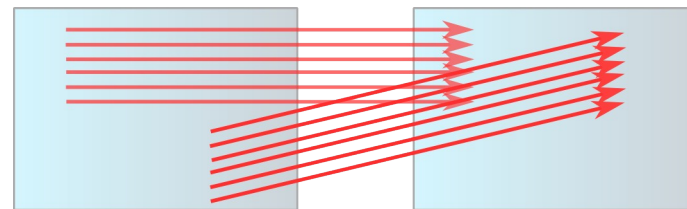
SIFT

- Feature-basiertes Verschiebungsfeld
- Features können überall im Raum liegen

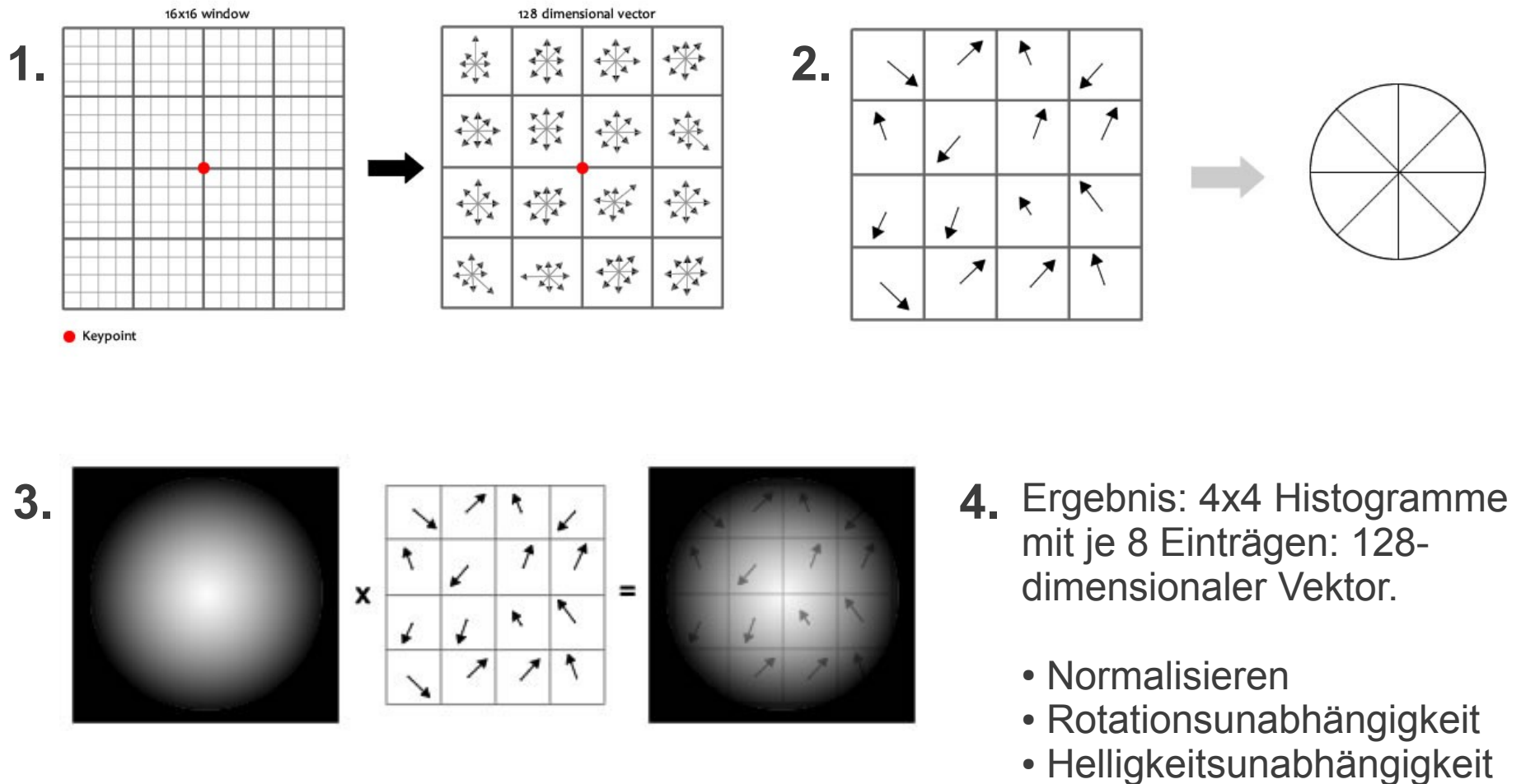


SIFT Flow

- Dichtes, feature-basiertes Verschiebungsfeld



SIFT Deskriptor



Algorithmus

Scene Matching

SIFT-Deskriptoren auswählen

SIFT-Deskriptoren gruppieren

Histogram-Intersection

Flow Algorithmus

SIFT-Deskriptor für jeden Pixel erstellen

Verschiebungsvektoren berechnen (Optimierungsproblem)

Algorithmus – Scene Matching

- Datenbank vorbereiten
 - 5000 zufällige SIFT-Deskriptoren auswählen
 - In 500 *visual words* einteilen mit K-means
- Histogramm der *visual words* erstellen mit einer zweistufigen *spatial pyramid*

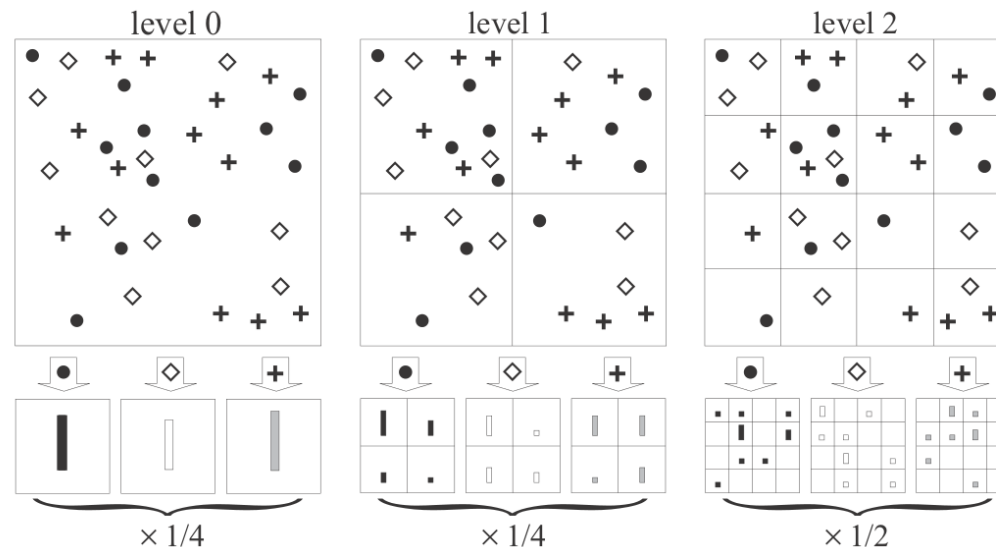
Scene Matching

SIFT-Deskriptoren auswählen

SIFT-Deskriptoren gruppieren

Histogramm-Intersection

spatial pyramid



Algorithmus – Scene Matching

- Mit *Histogram Intersection* die 20 ähnlichsten Bilder zu einem Referenzbild auswählen

Scene Matching

SIFT-Deskriptoren auswählen

SIFT-Deskriptoren gruppieren

Histogram-Intersection

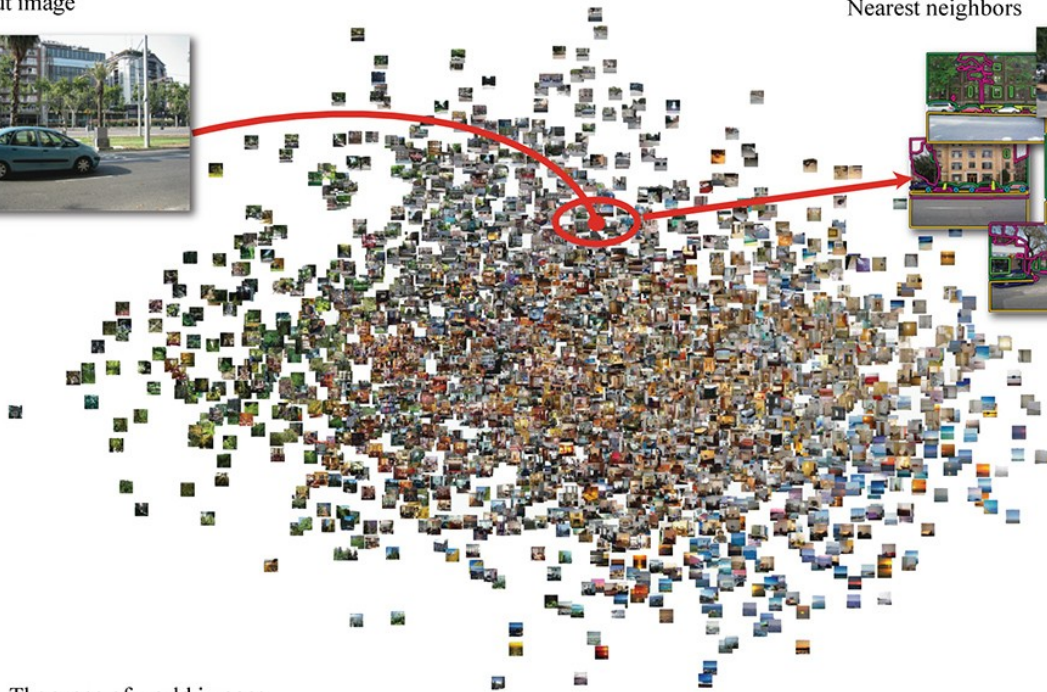
Input image



Nearest neighbors



The space of world images



Algorithmus – Scene Matching



Query image



Scene Matching

SIFT-Deskriptoren auswählen

SIFT-Deskriptoren gruppieren

Histogram-Intersection

Algorithmus - Flow

Flow Algorithmus

- Verschiebungsfeld berechnen mit zu minimierender Kostenfunktion

SIFT-Deskriptor für jeden Pixel erstellen

Verschiebungsfeld berechnen (Optimierungsproblem)

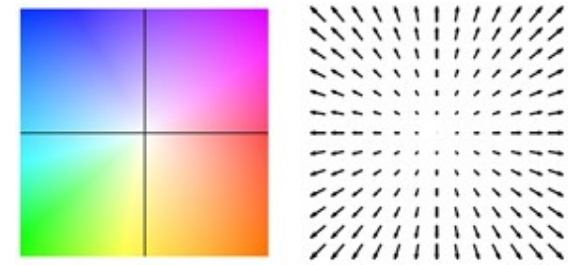
$$E(\mathbf{w}) = \sum_{\mathbf{p}} \|s_1(\mathbf{p}) - s_2(\mathbf{p} + \mathbf{w})\|_1 \quad \text{Abstand SIFT-Deskriptoren}$$
$$+ \frac{1}{\sigma^2} \sum_{\mathbf{p}} \left(u^2(\mathbf{p}) + v^2(\mathbf{p}) \right) \quad \text{Verschiebungsdistanz}$$
$$+ \sum_{(\mathbf{p}, \mathbf{q}) \in \varepsilon} \min \left(\alpha |u(\mathbf{p}) - u(\mathbf{q})|, d \right) + \min \left(\alpha |v(\mathbf{p}) - v(\mathbf{q})|, d \right) \quad \text{Glattheitsterme}$$

s : SIFT-Deskriptoren
 σ , α , d : Parameter

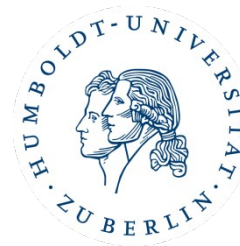
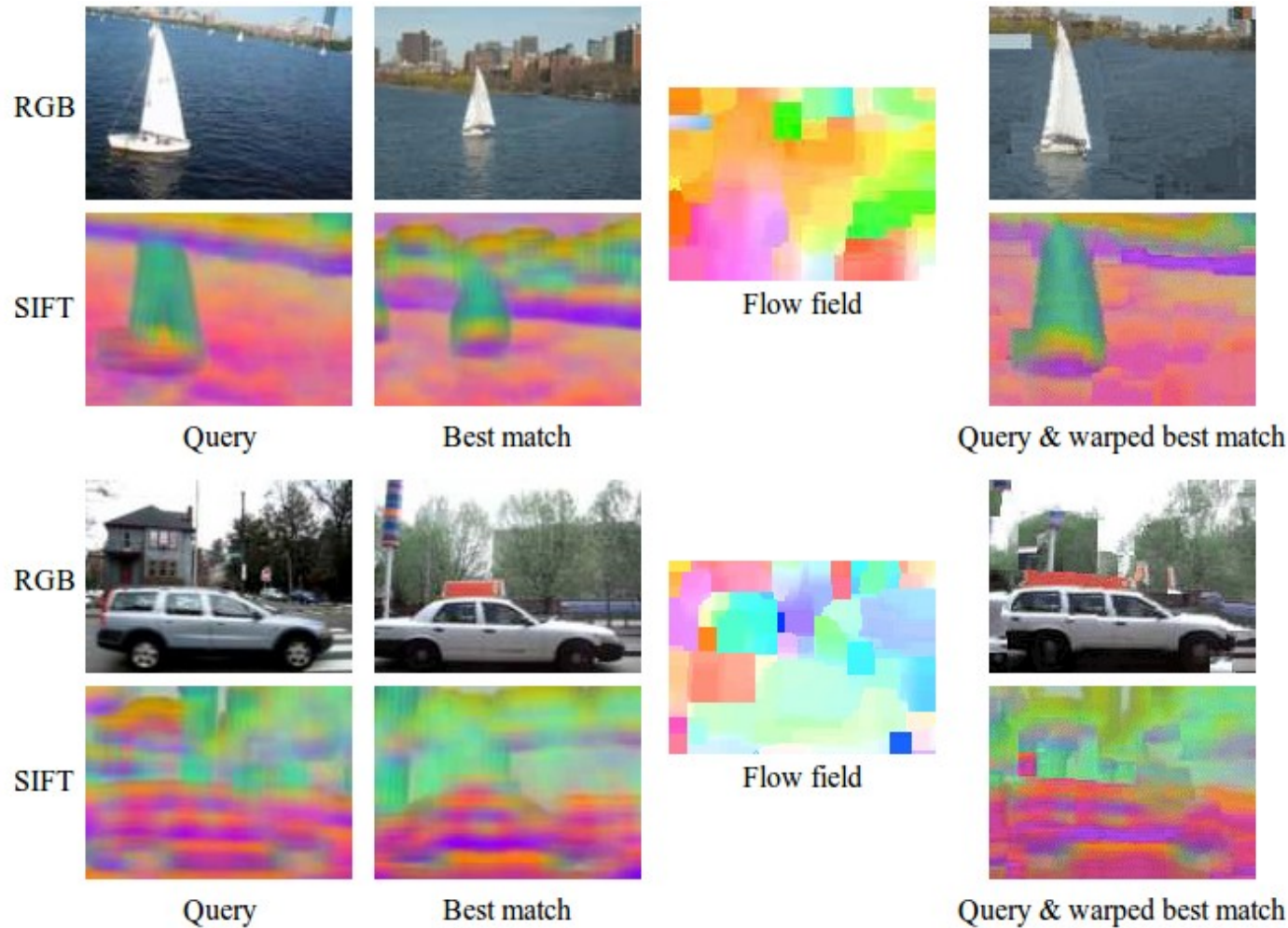
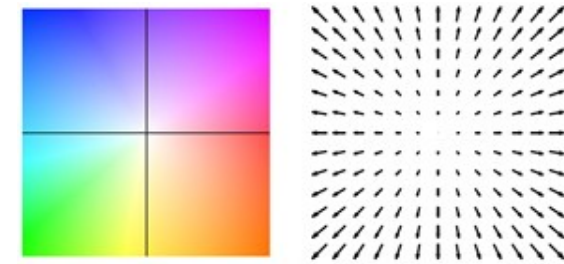
\mathbf{p} : $\mathbf{p}=(x,y)$, Pixelkoordinate
 \mathbf{w} : $\mathbf{w}(\mathbf{p})=(u(\mathbf{p}),v(\mathbf{p}))$, Verschiebungsvektor



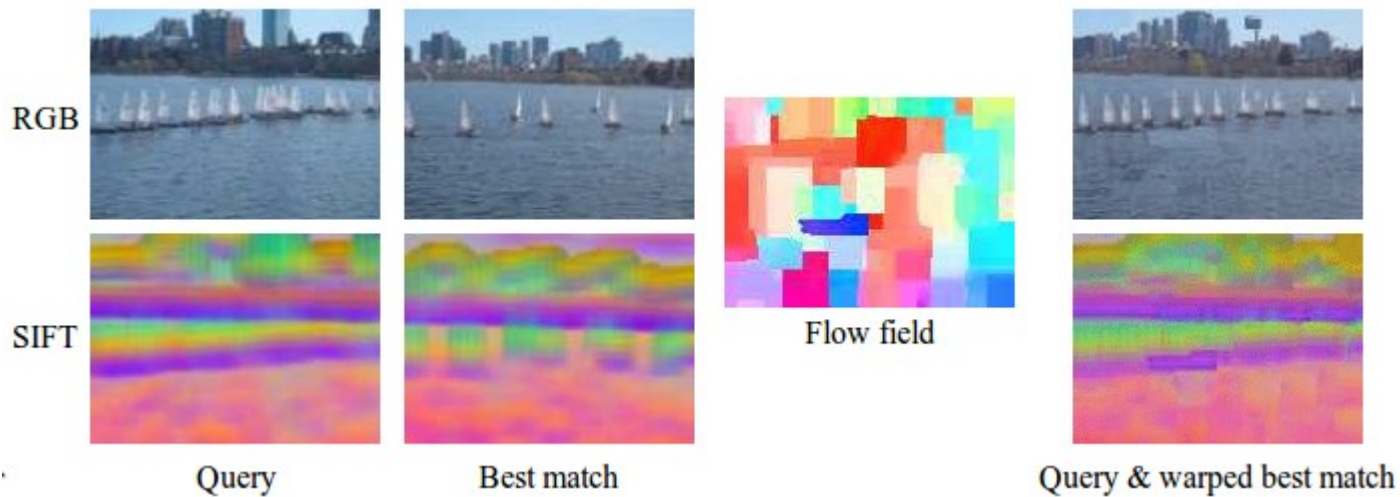
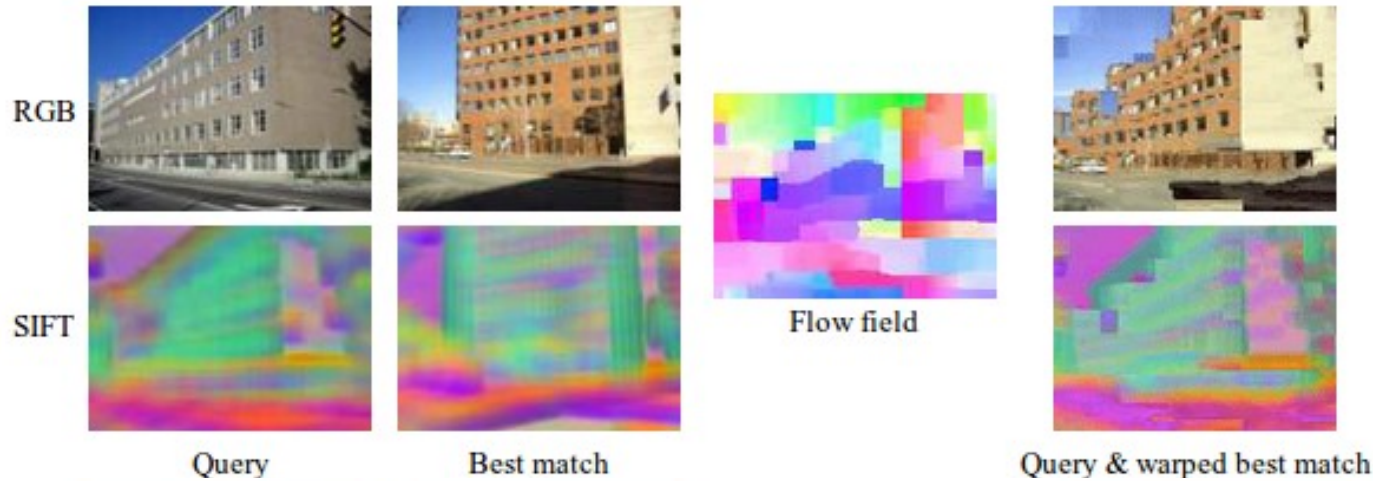
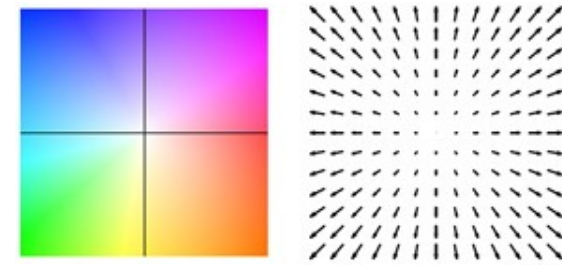
Ergebnisse



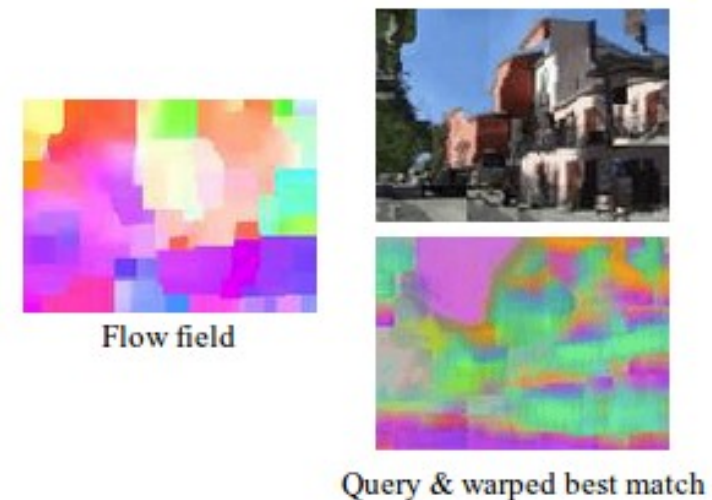
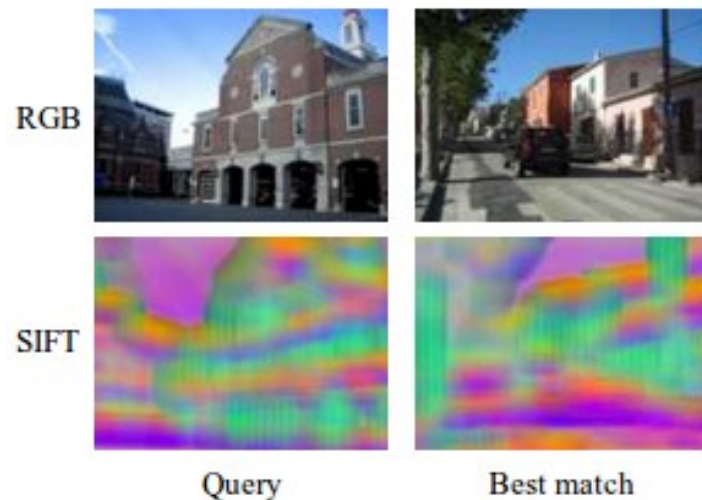
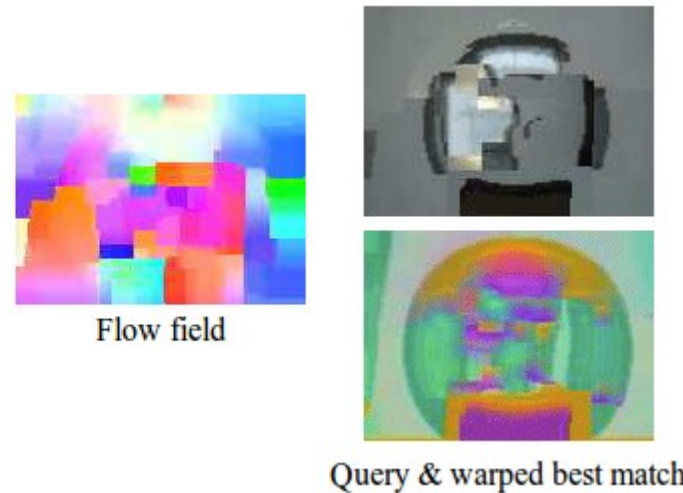
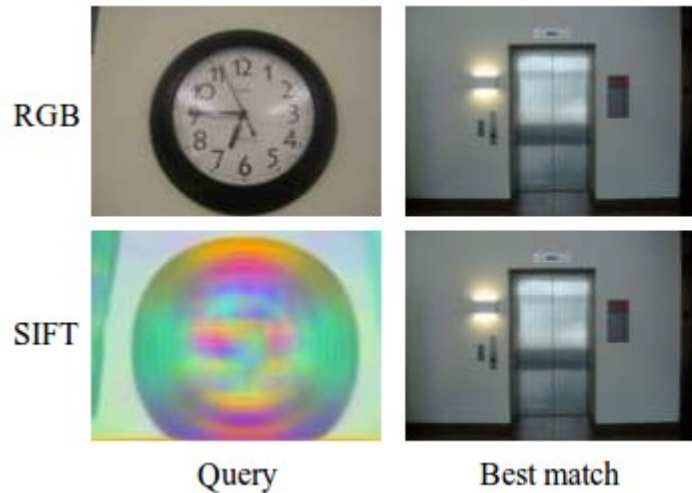
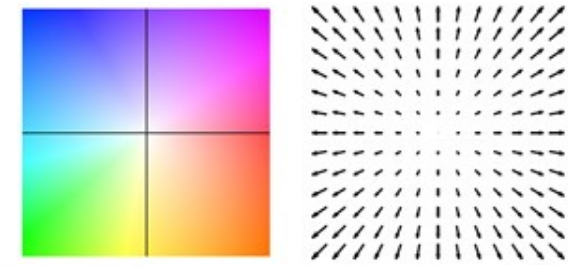
Ergebnisse



Ergebnisse



Ergebnisse



Anwendung – Motion Field Prediction

- Plausible Bewegungen eines einzelnen Bildes vorhersagen
- Datenbank mit Videos von Alltagsszenen
- Optischer Fluß zum Nachbarframe für jeden Frame gespeichert
- Suche zu einem Referenzbild ein ähnliches Bild aus der Datenbank (Scene Matching)
- Verschiebungsfeld mit dem Flow-Algorithmus berechnen
- Verschiebungsfeld auf den optischen Fluß anwenden



Anwendung – Motion Field Prediction



Original Image

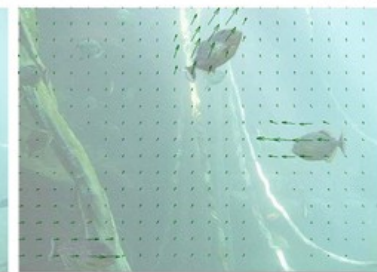
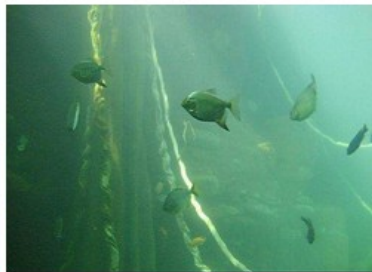
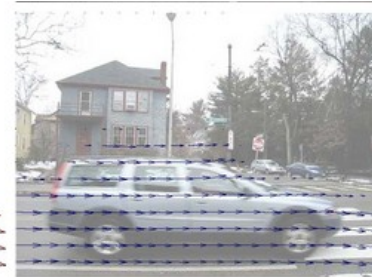
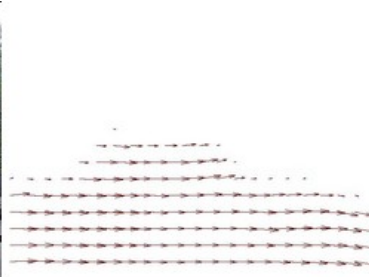
Database match

Motion of database
match

Warped and
transferred motion

Ground truth of
original

Anwendung – Motion Field Prediction



Original Image

Database match

Motion of database
match

Warped and
transferred motion

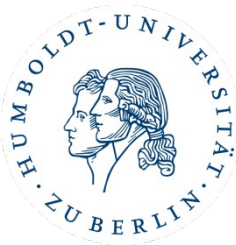
Ground truth of
original

Anwendung – Motion Field Prediction



Anwendung – Motion synthesis

- Bewegung in ein Standbild synthetisieren
- Bewegliche Objekte aus Videosequenzen extrahieren deren Szene ähnlich ist
- Scene Matching: zu einem Referenzbild eine ähnliche Videosequenz zu finden
- Für jeden Frame im Video die beweglichen Teile ermitteln
- Bewegliche Teile in das Referenzbild kopieren
- *Poisson Editing* um bewegliche Teile nahtlos zusammenzufügen



Anwendung – Motion synthesis



Anwendung – Motion synthesis



Quellen

- Liu, Yuen, Torralba, Sivic, Freeman: *SIFT Flow: Dense Correspondence across Different Scenes* (2008).
- Liu, Yuen, Torralba: *SIFT Flow: Dense Correspondence across Scenes and its Applications*. URL: <http://people.csail.mit.edu/celiu/SIFTflow/> - Zugriffsdatum: 23.6.2011.
- Lazebnik, Schmid, Ponce: *Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories*. In: Proc. CVPR, vol. II (2006).
- Baker, Scharstein, Lewis, Roth, Black, Szeliski: *A database and evaluation methodology for optical flow*. In: Proc. ICCV (2007).
- AI Shack: *SIFT: Scale Invariant Feature Transform*. URL: <http://www.aishack.in/2010/07/sift-step-6-generate-sift-features/> - Zugriffsdatum: 23.6.2011.

