

Erweiterung von Punktmerkmalen um Farbinformation zur semantischen Gruppierung

Nils Hering, Frank Schmitt, Lutz Prieße

Universität Koblenz-Landau,
Universitätsstr. 1, D-56070 Koblenz
E-Mail: nilshering, fschmitt, priese@uni-koblenz.de

Zusammenfassung Durch die semantische Gruppierung von Punktmerkmalen in Einzelbildern lassen sich Orte gleicher Bedeutung im Bild zusammenfassen. Dies kann in vielen Bereichen der Bildverarbeitung, wie z.B. der Objekterkennung, ein wichtiger initialer Schritt sein. Die gruppierten Merkmale können als Grundlage zu weiteren, spezielleren Untersuchungen des Bildes dienen, oder aber Ausgangspunkt für grob- granularere Deskriptoren sein. Bei vielen Typen von Punktmerkmalen (z.B. SIFT, vgl. [7]) wird die Information zum Aufbau des Deskriptors dabei unter Vernachlässigung von Farbe bestimmt. Andere Verfahren, die Punktmerkmale um Farbinformation ergänzen, zeigen sich aufgrund der Rechenintensität für die semantische Gruppierung oft ungeeignet oder zeigen Schwächen in bestimmten Anwendungsbereichen. In dieser Arbeit wird ein Verfahren zur Ergänzung von Punktmerkmalen, speziell SIFT-Merkmalen, um Farbinformation vorgestellt, welches die semantische Gruppierung selbiger in urbanen Umgebungen performant robuster macht.

1 Überblick

Die semantische Gruppierung von Punktmerkmalen wurde von den Autoren bereits in vorherigen Arbeiten untersucht (vgl. [4], [9]). Ziel der dort vorgestellten Verfahren ist es, alle SIFT-Merkmale in einem Bild, die Orte ähnlicher Semantik beschreiben, zu gruppieren. In [4] wurde dazu zunächst der Begriff der *Nachbarschaft* eingeführt. Unter der Nachbarschaft $N(f)$ eines Merkmals f in einem Bild I wird die Menge aller Merkmale f' in I verstanden, die nach einer Nachbarschaftsbeziehung zu f ähnlich sind. Die in [4] vorgestellte Nachbarschaftsbeziehung τ definiert die Ähnlichkeit zwischen Merkmalen anhand deren Ähnlichkeit in Skalierung, Orientierung und Beschreibungsvektor.

Mathematisch gesehen ist ein SIFT-Merkmal f ein Tupel $f = (s_f, o_f, v_f, l_f)$ bestehend aus vier Attributen: s_f für die Skalierung, o_f für die Orientierung, v_f für den 128-dimensionalen Beschreibungsvektor und l_f für den Ort des Merkmals im Bild in x- und

y-Koordinate. Der Euklidische Abstand $d_E(f_1, f_2)$ und die 7er-Distanz $d_7(f_1, f_2)$ zweier SIFT-Merkmale sind der Euklidische Abstand, bzw. die 7er-Distanz zwischen v_{f_1} und v_{f_2} . Dabei ist die 7er-Distanz als die Summe der sieben größten innerdimensionalen Abstände definiert. Der Wertebereich von o_f ist $[0, 2\pi[$, der von s_f hängt von der Bildgröße ab und liegt im hiesigen Anwendungsbereich etwa zwischen 0 und 100.

Nach der Nachbarschaftsbeziehung τ gehört $f' = (s', o', v', l')$ zu der (t_s, t_o, t_7) -Nachbarschaft von $f = (s, o, v, l)$, wenn gilt:

- $|s - s'| < t_s$,
- $|o - o'| < t_o$,
- $d_7(f, f') < t_7$

für drei Schwellwerte t_s, t_o, t_7 . Ein zugehöriger Algorithmus testet alle Merkmale auf deren potentielle Zugehörigkeit zu $N(f)$ entsprechend der o.g. Nachbarschaftsbeziehung, wobei er auch die mittleren Differenzen in Skalierung und Orientierung innerhalb der bereits gebildeten Nachbarschaft berücksichtigt.

Das Bilden der Nachbarschaften ist der initiale Schritt zum Zusammenfassen aller Bildorte mit gleicher Bedeutung. Dabei kann es jedoch zu zwei Problemen kommen: Oft decken die Nachbarschaften nicht den gesamten Bereich der Orte mit gemeinsamer Semantik ab, andererseits können auch Merkmale an Orten fremder Semantik Teil der Nachbarschaft sein.

Da alle Merkmale, die Teil einer Nachbarschaft sind, wiederum ihre eigene Nachbarschaft haben und diese auf Ähnlichkeit der Merkmale beruht, gibt es Nachbarschaften mit großen Schnittmengen untereinander. Diesen Sachverhalt kann man nutzen, um die Abdeckung der Orte mit gemeinsamer Semantik zu verbessern. Zur Vereinigung solcher Nachbarschaften wurde in [9] der *AGS*-Algorithmus (*A*utomatische *G*ruppierung von *S*emantik) eingeführt. AGS fasst ähnliche Nachbarschaften und bereits gebildete Gruppen zu s.g. *Mikrogruppen* (vgl. Abb. 1), zusammen. Ähnlichkeit ist dabei bzgl. der Überlappung von Orten von SIFT-Merkmalen zwischen den Gruppen definiert. Ist die Überlappungsrate $\frac{|loc(G) \cap loc(G')|}{\min(|G|, |G'|)}$ zwischen zwei Gruppen G und G' groß genug, gelten die Gruppen als ähnlich. Dadurch kann eine umfassendere Abdeckung der Orte ähnlicher Semantik im Bild erzielt werden.

Während die Gruppierung mittels *AGS* also bei der Lösung des Problems der unvollständigen Abdeckung von Semantiken hilft, sollen die Merkmale in dieser Arbeit um lokale Farbinformation erweitert werden, um beim Bilden der Nachbarschaften Verbesserungen beim Ausschluss semantikfremder Merkmale zu erhalten.

2 Ergänzung der Merkmale um Farbinformation

2.1 Einführung

In der Literatur findet man verschiedene Ansätze zur Kombination von SIFT-Deskriptoren mit Farbe. Abdel-Hakim und Farag [1] haben mit *CSIFT* einen neuen SIFT-verwandten farbinvarianten Deskriptor entwickelt. Der entscheidende Unterschied zur herkömmlichen SIFT-Version besteht darin, dass als Eingabebild die Farbinvarianten nach Geusebroek

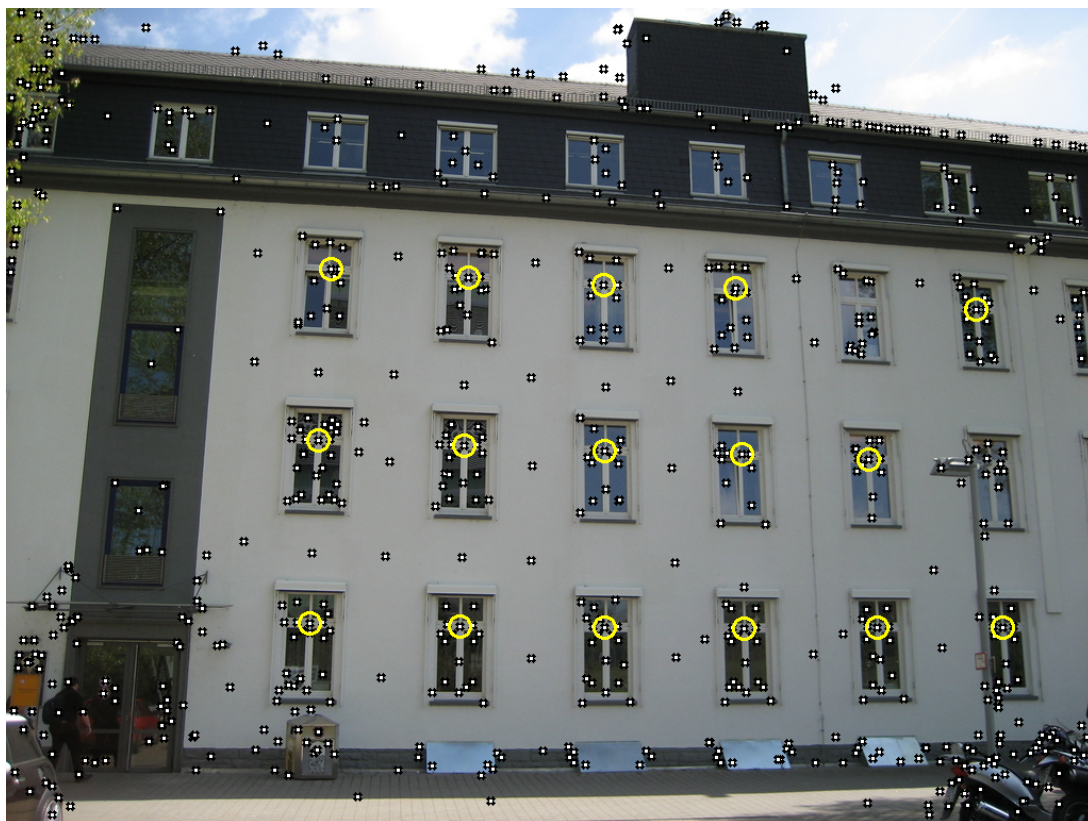


Abbildung 1: Die Abbildung zeigt die Mikrogruppe „Fensterkreuze“. Die schwarzweißen Kreuze markieren detektierte SIFT-Merkmale. Die farbigen Kringel umranden alle Merkmale der jeweiligen Mikrogruppe.

et al. [3] verwendet werden. Dem entsprechend werden die Deskriptoren nicht aus den Gradienten in den Grauwerten, sondern ebenfalls aus den Gradienten in den Farbinvarianten erzeugt. Hu *et al.* [5] nutzen den SIFT-Deskriptor kombiniert mit einem Farbdeskriptor zur Objekt-Wiedererkennung. Sie kombinieren SIFT dabei mit einem globalen Farbmerkmal, genannt *Color Autocorrelograms* (vgl. [6]). Van de Weijer *et al.* [11] kombinieren den SIFT-Deskriptor mit einem Farbwinkelhistogramm im HSV-Farbraum. Dabei wird das Histogramm in dem Bildbereich um das Merkmal herum berechnet, der auch zur Berechnung des SIFT-Merkmalsvektors anhand der darin vorkommenden Gradienten dient. Da Marcin Marszalek *et al.* [8] u.a. mit diesem lokalen Deskriptor nach van de Weijer *et al.* die PASCAL VOC Challenge 2007 gewannen, haben sich auch die Autoren zunächst an diesem Ansatz orientiert.

Diese Herangehensweise birgt jedoch ein Problem bzgl. des HSV-Farbraums und dem hier behandelten Anwendungsbereich. Im HSV-Farbraum existiert eine s.g. Unbuntachse entlang derer die Farben mit sehr geringer Sättigung angeordnet sind. Entlang dieser Achse ist der Farbwinkel undefiniert, bzw. stark schwankend. Dies kann wiederum in der Praxis zu großen Schwankungen zwischen der subjektiv empfundenen Ähnlichkeit zweier Farben und der Ähnlichkeit auf Datenebene bzgl. des Farbwinkels führen.

In Abb. 2 sieht man auf der linken Seite ein Bild eines Gebäudes des Universitätscampus in Koblenz. Dieses Bild ist typisch für den Anwendungsbereich der Algorithmen zur semantischen Gruppierung von Punktmerkmalen in urbanen Umgebungen. Auf der rechten Seite sieht man das gleiche Bild, allerdings überall außerhalb der glatten Fassadenwände

geschwärzt. Die Farbwerte der Pixel innerhalb der Fassade weisen im HSV-Farbraum eine sehr geringe Sättigung auf (einstelliger Prozentbereich). Dies führt zum o.g. Effekt der Undefiniertheit, bzw. der Schwankung der Farbwinkelwerte. Für den menschlichen Beobachter wirkt die Fassade sehr homogen, jedoch enthalten die Fassadenpixel auf Datenebene z.B. folgende Farbwinkelwerte (Wertebereich 0° - 360°):

- 0° (entspricht rot),
- 40° (orange),
- 72° (gelb-grün),
- 132° (grün),
- 165° (türkis),
- 220° (blau).



Abbildung 2: Ein herkömmliches Bild eines Gebäudes des Universitätscampus Koblenz, rechts nur die Fassadenwand ungeschwärzt

Das macht deutlich, dass die Farbwinkelwerte in gering gesättigten Farbbereichen (z.B. Gebäudefronten) stark schwanken und deshalb sinnvolle Vergleiche von Farbwinkelhistogrammen aus diesen Bereichen erschweren.

Eine mögliche Lösung für dieses Problem stellt z.B. die Verwendung des zweidimensionalen HS-Farbkreises dar. Da in einem durchschnittlichen Bild einer urbanen Umgebung der Auflösung 1024×768 aber rund 1000 SIFT-Merkmale detektiert werden, muss die Auswertung der Farbinformation in der Umgebung des Merkmals auch performant sein. Das spricht gegen einen 2-D-Ansatz und legt die Arbeit auf 1-D-Histogrammen nahe.

2.2 Histogrammberechnung

Die Grundidee dieser Arbeit ist es, zwischen bunten (Farbwinkel ist definiert) und unbunten (Farbwinkel kann undefiniert sein) Farben im HSV-Farbraum zu unterscheiden und nicht nur ein Farbwinkelhistogramm H^h , sondern auch ein Helligkeithistogramm H^v für unbunte Farben zu verwenden. Ob eine Farbe bunt oder unbunt ist, entscheidet man

anhand des Sättigungswertes. Ist also die Sättigung einer Farbe s kleiner als ein Schwellwert t_{sat} , gilt eine Farbe als unbunt. In diesem Fall entscheidet ihr Helligkeitswert v , in welchen Bin die Farbe im Helligkeitshistogramm H^v einsortiert wird. Gilt $s > t_{sat}$, wird die Farbe anhand ihres Farbwinkels h in das Farbwinkelhistogramm H^h einsortiert. In den Experimenten zu dieser Arbeit wurde ein t_{sat} -Wert von ca 23% gewählt. Formal gesehen ordnet also eine Funktion ξ jedem HSV-Farbpixel $f_{ij} = (h_{ij}, s_{ij}, v_{ij})$ eine weitere Funktion zu, welche das Farbpixel entweder in $H^v(\xi^v)$ oder in $H^h(\xi^h)$ einordnet.

$$\xi(f_{ij}) : \begin{cases} \xi^h(f_{ij}) & , \text{ falls } s_{ij} < t_{sat} \\ \xi^v(f_{ij}) & \text{ sonst} \end{cases}$$

Die beiden Funktionen ξ^h und ξ^v sind am Beispiel von ξ^h folgendermaßen definiert (ξ^v analog):

$$\xi^h(f_{ij}) : \mathbb{N}^3 \rightarrow \{1, \dots, N_h\},$$

$$\xi^h(f_{ij}) = \lceil \frac{h_{ij}}{\lceil \frac{h_{max}}{N_h} \rceil} \rceil,$$

wobei N_h für die Anzahl der Bins in Histogramm H^h steht und h_{max} den maximalen Wert meint, den h annehmen kann. Je nach Wertebereich und Anzahl der Bins kann es dabei zu Randfällen kommen, die speziell betrachtet werden müssen. Bei den angeführten Histogrammen handelt es sich generell um relative Histogramme.

Zunächst gilt es den Bildbereich für die Histogrammberechnung pro Merkmal festzulegen, d.h. zu bestimmen, welcher Bildausschnitt in der Nachbarschaft des Merkmals die Eingabe für die Histogrammberechnung darstellt. Dazu hat man generell zwei Möglichkeiten. Zum einen kann man den Bildbereich entsprechend der Skalierung des Merkmals adaptieren. Somit würde sich der einbezogene Farbbildbereich entsprechend dem einbezogenen Gradientenbildbereich von Merkmal zu Merkmal ändern. Eine andere Möglichkeit besteht darin, den Bildbereich fix für jedes Merkmal zu wählen.

In dieser Arbeit wurde der Bildbereich zur Erzeugung der Histogramme, ähnlich dem Einzugsbereich für die Bildung des Gradientenvektors bei SIFT, dynamisch gewählt. Die Seitenlänge des quadratischen Bildausschnitts berechnet sich dabei auf Grundlage der Skalierung des betrachteten Merkmals. Es wird dazu die gleiche Formel wie im SIFT-Algorithmus (vgl. [7]) verwendet, allerdings ist die Seitenlänge auf ca 10% der Bildbreite beschränkt. Übersteigt die Länge der Seiten diese 10%, wird sie auf eben diesen Wert neu festgesetzt.

Ein 1-D-Grauwertshistogramm über ein gesamtes Bild wird normalisiert, indem man die absolute Häufigkeit jedes Grauwertes durch die Gesamtzahl aller Pixel im Bild dividiert. In der Variante mit zwei 1-D-Histogrammen werden aber zwei Histogramme aus nur einem Bild(-bereich) erzeugt. Die beiden Histogramme sind dabei disjunkt, d.h. jedes Pixel, bzw. dessen Farbwert, kann nur in genau ein Histogramm eingeordnet werden. Die Summe der Summen der absoluten Häufigkeiten der jeweiligen Histogrammbins aus beiden Histogrammen ist also gleich der Gesamtzahl der im Bild(-bereich) vorkommenden Pixel, bzw. Farbwerte.

Schließlich können die Histogramme auf zwei Arten normalisiert werden: Entweder, indem die absolute Häufigkeit jedes Bins durch die Gesamtzahl der im Bild vorkommenden Pixel dividiert wird (vereinigt), oder aber indem die absolute Häufigkeit jedes Bins nur durch

die Gesamtzahl der in das jeweilige Histogramm eingeordneten Farbwerte dividiert wird (separiert) (vgl. Abb. 3).

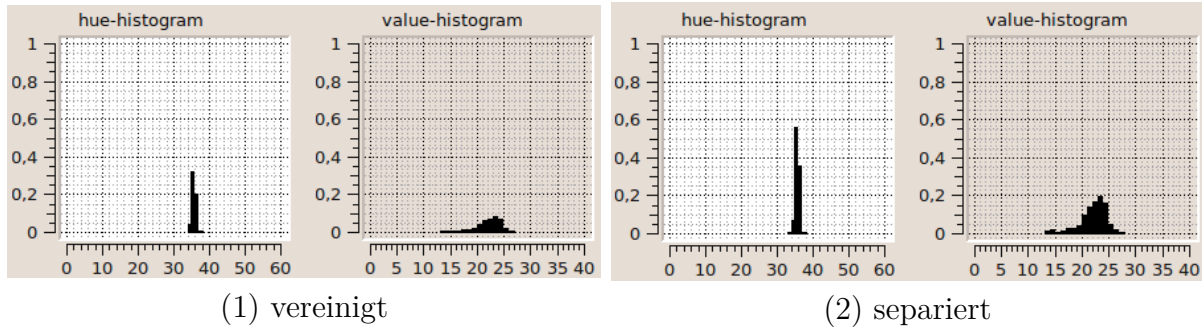


Abbildung 3: Gegenüberstellung von vereinigten (links) und separierten Histogrammen (rechts)

Hier wurde die vereinigte Variante (vgl. Abb. 3, links (1)) gewählt. Das Verhältnis der Pixelmengen zwischen den Histogrammen bleibt so erhalten.

2.3 Distanzberechnung zwischen Histogrammen

In der Nachbarschaftsbeziehung auf SIFT-Merkmalen (vgl. Abschnitt 1) wurde festgelegt, wann ein SIFT-Merkmal zu einer Nachbarschaft eines anderen SIFT-Merkmals gehört. Diese Nachbarschaftsbeziehung muss nun an die neuen, um Farbhistogramme ergänzten, Deskriptoren angepasst werden.

Nach der neuen Nachbarschaftsbeziehung τ_c gehört das Merkmal $f' = (s', o', v', l', H^{v'}, H^{h'})$ zu der (t_s, t_o, t_7, t_H) -Nachbarschaft von $f = (s, o, v, l, H^v, H^h)$, wenn gilt:

- $|s - s'| < t_s$,
- $|o - o'| < t_o$,
- $d_7(f, f') < t_7$
- $d_H(f, f') < t_H$

Zusätzlich zu den bekannten Schwellwerten t_s , t_o und t_7 kommt jetzt also noch der Schwellwert t_H hinzu. Um t_H sinnvoll angeben zu können, muss zunächst $d_H(f, f')$ definiert werden. Grob gesprochen ist $d_H(f, f')$ ein Distanzmaß zwischen den beiden Histogrammen des Merkmals f und denen des Merkmals f' . Um dieses Distanzmaß passend zu wählen, muss man zunächst auf 1-D-Histogrammen definierte Abstandsmaße auf ihre Tauglichkeit für den konkreten Anwendungsbereich prüfen.

Rubner *et al.* bieten in [10] eine gute Zusammenstellung von 1-D-Histogrammabstandsmaßen. Dort wird vor allem zwischen „Bin-By-Bin Dissimilarity Measures“, also Abstandsmaßen die auf Bin-zu-Bin Vergleichen basieren, und „Cross-Bin Dissimilarity Measures“, also komplizierteren Abstandsmaßen, die z.B. auf kumulative Histogramme oder zweidimensionale Bin-zu-Bin Abstandsmatrizen zurückgreifen, unterschieden. Als Beispiele für „Bin-By-Bin Dissimilarity Measures“ gibt es u.a.:

Minkowski – FormDistance

$$d_{L_r}(H, K) = \left(\sum_i |h_i - k_i|^r \right)^{1/r},$$

Sum Of Squared Distances(SSD)

$$d_{SSD}(H, K) = \left(\sum_i |h_i - k_i|^2 \right),$$

Nach Rubner *et al.* haben diese einfachen Methoden jedoch einige Nachteile. Danach liegt der Hauptnachteil in Ihrer Eigenart, nur die Übereinstimmung zwischen Bins mit den selben Indizes zu messen. Weiterhin ist die Empfindlichkeit gegenüber der Bin-Größe nach Rubner ein Problem. Zu grobe Quantisierungen können demnach nicht stark genug unterscheiden und bei einer zu feinen Quantisierung fallen u.U. sehr ähnliche Werte in unterschiedliche Bins und können somit nicht verglichen werden.

Um diese Nachteile zu umgehen, kann man die bereits erwähnten „Cross-Bin Dissimilarity Measures“ verwenden, z.B. die *Match – Distanz*, welche folgendermaßen definiert ist:

Match – Distanz

$$d_M(H, K) = \sum_i |\hat{h}_i - \hat{k}_i|,$$

wobei $\hat{h}_i = \sum_{j \leq i} h_j$ das *kumulative Histogramm* von $\{h_i\}$ ist. Analoges gilt für $\{k_i\}$. Im Rahmen dieser Arbeit wurden Minkowski-Form-Distanz, SSD und die Match-Distanz näher getestet. Trotz der Einfachheit der Methode wurde aufgrund der besseren Ergebnisse letztendlich die Minkowski-Form-Distanz mit $r = 1$ ausgewählt.

Im sich hier darstellenden Anwendungsfall von zwei 1-D-Histogrammen kommt ein Problem bei der Abstandsmessung hinzu, denn alle hier genannten Maße sind nur auf einem Histogramm definiert. Da die Nachbarschaftsalgorithmen ebenso nur einen zusätzlichen Schwellwert für die Nachbarschaftsbildung erhalten sollen, muss man die beiden Abstandswerte $d_{H^v}(f, f')$ und $d_{H^h}(f, f')$ kombinieren. Hier wurde definiert: $d_H(f, f') = d_{H^v}(f, f') + d_{H^h}(f, f')$. Enthält das Histogrammpaar eines Merkmals f überwiegend bunte Farben (mehr als $\frac{2}{3}$ aller Farbwerte) und ein Paar eines anderen Merkmals f' überwiegend unbunte (ebenso mehr als $\frac{2}{3}$ aller Farbwerte), so wird der Histogrammabstand $d_H(f, f')$ zwischen beiden als unendlich angenommen.

3 Evaluation

Zur Evaluation der vorgestellten Erweiterung von Punktmerkmalen um Farbinformation wurden Mikrogruppen verschiedener Varianten von SIFT-Merkmalen gegenübergestellt. Auf der einen Seite stehen dabei Mikrogruppen auf Basis von SIFT_{noi}-Merkmalen, einer Variante von SIFT-Merkmalen bei denen die Beschreibungsvektoren nicht anhand der Orientierung normalisiert werden und die Orientierungsunterschiede der Merkmale bei der Nachbarschaftstopologie nicht betrachtet werden (vgl. [9]). Diesen werden Mikrogruppen auf Basis von um Farbinformation erweiterten SIFT_{noi}-Merkmalen gegenübergestellt. Beispiele für solche Bilder sind im Anhang zu finden.

Untersucht wurden Mikrogruppen aus 75 Bildern, welche wiederum zufällig aus einer Datenbasis von 300 Bildern des Universitätscampus Koblenz gewählt wurden. Es wurden

zunächst diejenigen Mikrogruppen selektiert, bei denen im SIFT_{noi} -Fall ohne Farbinformation beurteilt werden konnte, ob die sie Fehler enthalten oder nicht. Aus diesen Gruppen wurden wiederum solche ausgesucht, bei denen im SIFT_{noi} -Fall ohne Farbinformation augenscheinlich Fehler vorhanden sind, die bei sinnvoller Betrachtung von Farbinformation zu vermeiden wären. Das bedeutet, dass die fehldetektierten Merkmale in deutlich anderen, semantikfremden Farbbereichen liegen. Bei diesen Gruppen wurde über die Maße *Coverability Rate* (*Cov*) und *Error Rate* (*Err*) bewertet wie gut die Fehler durch die Betrachtung von Farbe eliminiert werden konnten und wie stark dadurch die Abdeckung der korrekten Bildorte in Mitleidenschaft gezogen wurden (vgl. Tab. 1).

Versteht man unter G die erzeugte Mikrogruppe und unter GT eine „Ground Truth“ von Orten tatsächlicher semantischer Zusammengehörigkeit die G möglichst gut abdecken soll, dann sind *Coverability Rate* $\text{Cov}(G, GT)$ und *Error Rate* $\text{Err}(G, GT)$ folgendermaßen definiert:

$$\text{Cov}(G, GT) := \frac{|\text{loc}(G) \cap GT|}{|GT|}$$

$$\text{Err}(G, GT) := \frac{|\text{loc}(G) - GT|}{|G|}$$

$\text{loc}(G) := \{\text{loc}(f) | f \in G\}$ beschreibt dabei die Orte der gebildeten Mikrogruppen, wobei $\text{loc}(f)$ die x- und y-Koordinate des Merkmals f im Bild beschreibt.

Mikrogruppen	mittlere Cov	mittlere Err
112	84.8	27.3

Mikrogruppen	mittlere Cov	mittlere Err
112	83.5	5.3

Tabelle 1: Gegenüberstellung von mittlerer *Coverability Rate* und mittlerer *Error Rate* von fehlerhaften Mikrogruppen auf SIFT_{noi} -Merkmalen ohne Farbinformation (oben) und mit Farbinformation (unten).

Die Ergebnisse machen deutlich, dass durch die Ergänzung der Merkmale um Farbinformation eine deutliche Reduzierung der *Err* bei nur minimaler Reduzierung der *Cov* gelungen ist. Dabei müssen jedoch die Randbedingungen beachtet werden. 651 der 828 betrachteten Gruppen (rund 79%) enthielten keine Fehler. Unter den restlichen 177 Gruppen befanden sich 112 Gruppen (rund 63%) deren Fehler in semantikfremden Farbbereichen lagen. Das bedeutet, dass die positiven Effekte der Betrachtung von Farbe nur auf etwa 14% der Mikrogruppen Auswirkungen haben. Die Zeit zur Berechnung der Merkmale und deren Abstände untereinander steigt dagegen im Schnitt um 17%. Zur Durchführung der Experimente ist ebenfalls zu beachten, dass die Beurteilung der Qualität einer Mikrogruppe immer eine subjektive Komponente enthält. Beispiele zu Paaren von SIFT_{noi} -Gruppen ohne und mit Farbbetrachtung sind im Anhang zu finden.

4 Fazit

Es wurde ein Verfahren zur Ergänzung von Punktmerkmalen um Farbinformation vorgestellt. Dieses Verfahren ist speziell an die Anforderungen der Bildung von Mikrogruppen

in urbanen Umgebungen angepasst und liefert dort zuverlässige Ergebnisse. In einer Evaluation konnte gezeigt werden, dass es, mittels der in der Arbeit vorgestellten Methode zur Erweiterung von Punktmerkmalen um Farbinformation, möglich ist, die Qualität der Mikrogruppen mit Fehldetektionen in semantikfremden Farbbereichen deutlich zu verbessern. Es bleibt aber festzuhalten, dass die vorgestellte Methode nur auf etwa 15 % der untersuchten Mikrogruppen sinnvoll anzuwenden ist. Außerhalb urbaner Umgebungen sind die Methoden zur semantischen Gruppierung von Punktmerkmalen allgemein und speziell die vorgestellte Methode ihren Nutzen bisher schuldig geblieben.

Literatur

- [1] ABDEL-HAKIM, ALAA E. und ALY A. FARAG: *CSIFT: A SIFT Descriptor with Color Invariant Characteristics*. Computer Vision and Pattern Recognition, IEEE Computer Society Conference on, 2:1978–1983, 2006.
- [2] BALTHASAR, DIRK: *Drei neue Verfahren zum Matching und zur Klassifikation unter Echtzeitbedingungen*. Doktorarbeit, Universität Koblenz, Verlag Foelbach Koblenz, 2006.
- [3] GEUSEBROEK, J.M., R. VAN DEN BOOMGAARD, A.W.M. SMEULDERS und H. GEERTS: *Color Invariance*. PAMI, 23(12):1338–1350, December 2001.
- [4] HERING, NILS, FRANK SCHMITT und LUTZ PRIESE: *Image Understanding Using Self-similar SIFT features*. In: *Fourth International Conference on Computer Vision Theory and Applications (VISAPP) 2009, Lisboa, Portugal*, Band 2, Seiten 114–119, 2009.
- [5] HU, L., S.Q.A. JIANG, Q.M. HUANG und W. GAO: *People re-detection using Adaboost with sift and color correlogram*. In: *ICIP08*, Seiten 1348–1351, 2008.
- [6] HUANG, J., S.R. KUMAR, M. MITRA, W.J. ZHU und R. ZABIH: *Image Indexing Using Color Correlograms*. In: *CVPR97*, Seiten 762–768, 1997.
- [7] LOWE, DAVID G.: *Distinctive Image Features from Scale-Invariant Keypoints*. International Journal of Computer Vision, 60(2):91–110, 2004.
- [8] MARSZALEK, MARCIN, CORDELIA SCHMID, HEDI HARZALLAH und JOOST VAN DE WEIJER: *Learning Object Representations for Visual Object Class Recognition*, oct 2007. Visual Recognition Challenge workshop, in conjunction with ICCV.
- [9] PRIESE, LUTZ, FRANK SCHMITT und NILS HERING: *Grouping of Semantically Similar Image Positions*. In: SALBERG, ARNT-BORRE, JON YNGVE HARDEBERG und ROBERT JENSSEN (Herausgeber): *16th Scandinavian Conference, SCIA 2009, Oslo, Norway, June 15-18, Proceedings*, Band 5575 der Reihe *Lecture Notes in Computer Science*, Seiten 726–734, 2009.
- [10] RUBNER, Y., C. TOMASI und L.J. GUIBAS: *The Earth Mover's Distance as a Metric for Image Retrieval*. IJCV, 40(2):99–121, November 2000.

- [11] WEIJER, JOOST VAN DE, THEO GEVERS und ANDREW D. BAGDANOV: *Boosting Color Saliency in Image Feature Detection*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 28(1):150–156, 2006.

Beispielbilder



Abbildung 4: Im Bild oben ist eine Mikrogruppe aus SIFT_{noi} -Merkmalen ohne Farbinformation zu sehen. Man erkennt deutlich das semantikfremde Merkmal an der Rasengrenze. Im Bild unten sieht man die gleiche Mikrogruppe auf Basis von SIFT_{noi} -Merkmalen mit Farbinformation. Hier wird das semantikfremde Merkmal korrekt ausgeschlossen. Die schwarzweißen Kreuze markieren detektierte SIFT_{noi} -Merkmale. Die farbigen Krinkel umranden alle Merkmale der jeweiligen Mikrogruppe.

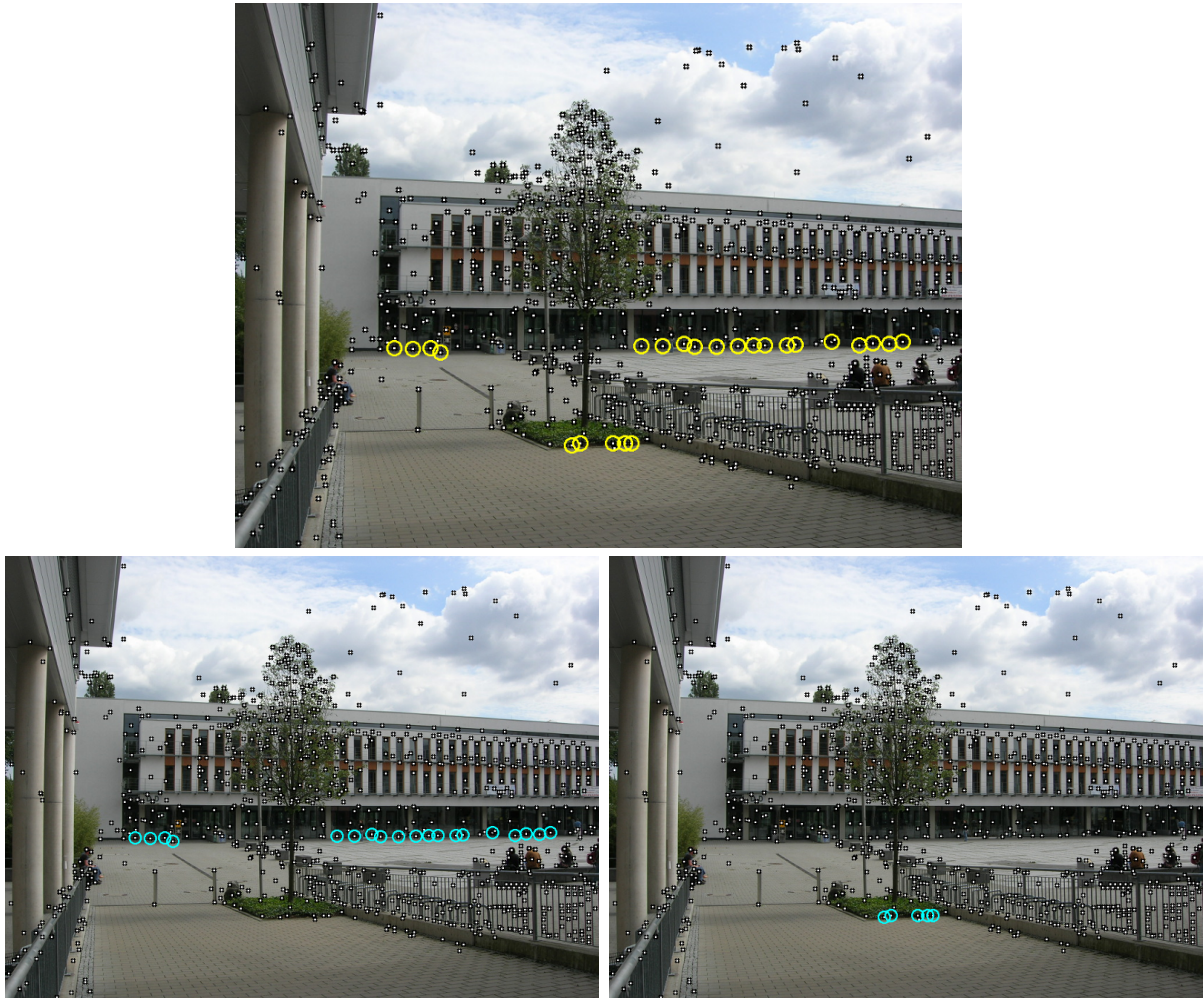


Abbildung 5: Im Bild oben ist erneut eine Mikrogruppe auf Basis von SIFT_{noi} -Merkmalen ohne Farbinformation zu sehen. Die Bilder unten links und rechts zeigen zwei Mikrogruppen auf Basis von SIFT_{noi} -Merkmalen mit Farbinformation. Es wird deutlich, dass die Mikrogruppe aus dem oberen Bild korrekterweise aufgrund der Farbinformation in zwei separate Gruppen getrennt wurde.