

Detektion wiederholter und symmetrischer Strukturen in Fassadenbildern

SUSANNE WENZEL, MARTIN DRAUSCHKE, WOLFGANG FÖRSTNER¹

Zusammenfassung:

Regelmäßige Strukturen und Symmetrien kennzeichnen viele Gebäudefassaden oder Objekte im Umfeld von Gebäuden. Für die automatisierte Bildinterpretation weisen diese Strukturen auf künstliche Objekte hin, führen aber auch zu Schwierigkeiten bei klassischen Bildzuordnungsverfahren. Die Suche und Gruppierung zusammengehöriger Merkmale kann daher sowohl zur Identifikation künstlicher Objekte als auch zur Verbesserung von Zuordnungsverfahren dienen.

Für die Analyse von entzerrten Fassadenaufnahmen haben wir das Verfahren von [LOY 2006] zur Detektion symmetrischer Bildstrukturen zu einem Verfahren zur Detektion verschiedener, sich wiederholender Bildstrukturen erweitert und aus den detektierten wiederholten Objekten eine minimale Beschreibung der Struktur der Fassadenelemente in Form von achsenparallelen Basiselementen abgeleitet.

1 Einleitung

Viele Gebäude werden durch symmetrische bzw. sich wiederholende Strukturen gekennzeichnet, insbesondere ist die Anordnung der Fenster in einer Fassade oftmals durch ein festes Muster geprägt. Diese Regelmäßigkeiten weisen auf künstliche Objekte hin, die wir für die automatisierte Bildinterpretation nutzen wollen. Die Suche und Gruppierung zusammengehöriger Merkmale kann daher sowohl zur Identifikation künstlicher Objekte als auch zur Verbesserung von Zuordnungsverfahren dienen, wo wiederholte Strukturen im Allgemeinen Schwierigkeiten bereiten.

Wir gehen davon aus, dass sich wiederholende Strukturen in Fassaden in horizontaler und vertikaler Richtung verlaufen. Aus diesem Grund entzerren wir die Bilder manuell vor der Merkmalsextraktion. Dieser Schritt könnte auch automatisiert werden, indem die beiden wichtigsten Fluchtpunkte nach [Schaffalitzky 2000] bestimmt werden.

Die Ableitung des Musters der Fassadenelemente ist für das modellbasierte Lernen im Kontext der Gebäudeerkennung von hoher Bedeutung. Des Weiteren kann über die Struktur des Musters auf Lücken in der Merkmalsextraktion geschlossen werden. Das Ziel dieser Arbeit ist die Bestimmung einer minimalen Beschreibung der Struktur der Fassadenelemente in Form von achsenparallelen Basiselementen. Diese sollen die Verteilung der Grundelemente der Fassade, insbesondere der Fenster, in Form von Translationsvektoren und der Anzahl der auftretenden Wiederholungen der Objekte in der entsprechenden Richtung charakterisieren.

¹ Rheinische Friedrich-Wilhelms-Universität Bonn, Institut für Geodäsie und Geoinformation, Professur für Photogrammetrie, Nussallee 15, 53115 Bonn; {susanne.wenzel,martin.drauschke}@uni-bonn.de, wf@ipb.uni-bonn.de

2 Symmetriedetektion

Das hier vorgestellte Verfahren basiert auf der Arbeit von [LOY 2006], in der ein Verfahren zur Detektion dominanter Symmetrien vorgestellt wurde. Die grundlegende Vorgehensweise ist in Abbildung 1 skizziert. Zunächst werden markante Bildpunkte detektiert. Wir verwenden den SIFT-Operator nach [LOWE 2004], da bei diesem Verfahren jeder extrahierte Punkt durch seine geometrischen Eigenschaften (Zeile, Spalte, Maßstab und Orientierung) sowie

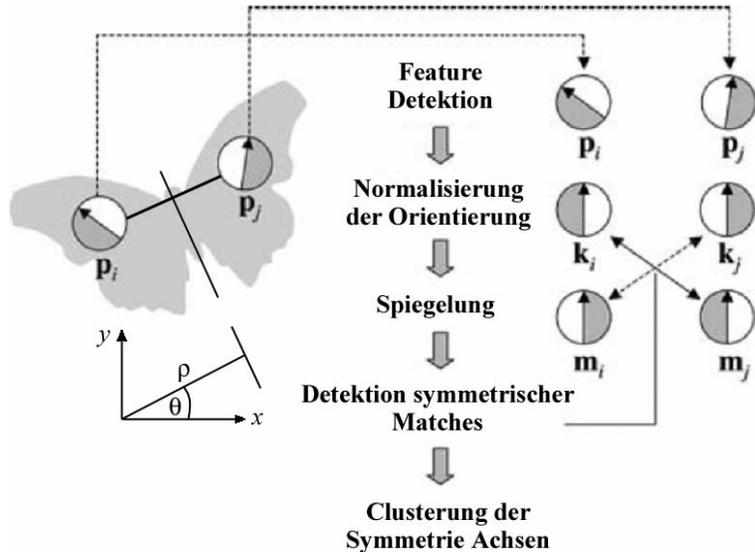


Abbildung 1: Schematische Darstellung der Detektion symmetrischer Merkmals-Paare, nach [LOY 2006].

Abbildung 1: Schematische Darstellung der Detektion symmetrischer Merkmals-Paare, nach [LOY 2006].

einen so genannten Deskriptor beschrieben wird, der die lokale Umgebung des Merkmals geeignet codiert. Bei Verwendung der SIFT-Merkmale sind diese Deskriptoren in Richtung der Orientierung des Merkmals normiert und damit rotationsinvariant.

Durch Umsortierung der Elemente des Deskriptor-Vektors können gespiegelten Versionen der detektierten Merkmale sehr einfach abgeleitet werden [Wenzel 2006].

Zuordnung des originalen Merkmals-Satzes zu den gespiegelten Kopien führt zu Paaren potentiell symmetrischer Merkmale. Die Symmetrieachse jedes Paares wird durch ihre Orientierung θ und ihren Abstand ρ zum Ursprung repräsentiert.

Die im Bild dominanten Symmetrien können dann durch Clustering der gefundenen Symmetrieachsen abgeleitet werden. Zur Bewertung der Qualität der Symmetrie der zugeordneten Merkmals-Paare wird eine Gewichtung der Symmetrie in Abhängigkeit von Orientierung, Skalendifferenz und Distanz der Merkmale eingeführt.

2.1 Bewertung der Symmetrie einzelner Merkmals-Paare

Mit dem Winkel-Symmetrie-Gewicht $\Phi_{ij} \in [-1,1]$

$$\Phi_{ij} = -\cos(\varphi_i + \varphi_j - 2\theta) = -\cos(\alpha + \beta) \quad (1)$$

werden diejenigen Merkmals-Paare hoch bewertet, deren Orientierung möglichst symmetrisch bzgl. der ermittelten Symmetrieachse dieses Paares ist. Abbildung 2 verdeutlicht den Einfluss der Orientierungen der Merkmale bzgl. ihrer Verbindungslinie. Die Winkel α und β ergänzen sich zu 180° , falls die Orientierungen der Merkmale exakt spiegelsymmetrisch zur Symmetrieachse der Merkmale sind.

Zwei weitere Gewichte gehen in die Bewertung von Symmetrien ein, das Skalengewicht S_{ij} sowie das Distanzgewicht

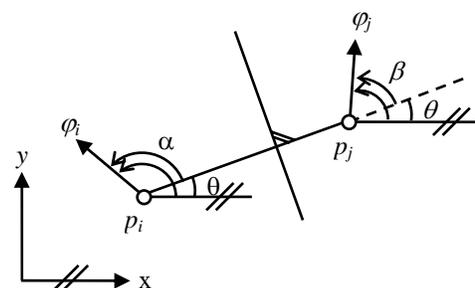


Abbildung 2: Bewertung der Symmetrie der Orientierung der Merkmale p_i und p_j . Wenn die Merkmale genau spiegelsymmetrisch orientiert sind, ergänzen sich die Winkel α und β , zwischen Hauptorientierung und Verbindungslinie der Merkmale zu 180° .

D_{ij} , siehe dazu [LOY 2006]. Das Skalengewicht beurteilt die Maßstabsdifferenzen zwischen den Merkmals-Paaren. Für Merkmalspunkte aus weit auseinander liegenden Skalen liegt S_{ij} nahe bei 0, bei identischen Skalen nimmt S_{ij} den Wert 1 an. Durch das Distanzgewicht kann der Abstand zur potentiellen Symmetrieachse und damit die Ausdehnung der symmetrischen Objekte überprüft werden. Liegen zwei Merkmalspunkte weit von einander entfernt, dann ist D_{ij} sehr klein, bei identischen Punkten wiederum 1.

Die Bewertung der Qualität der Symmetrie zweier Merkmale p_i und p_j ergibt sich schließlich aus allen Gewichten als Symmetriegrad M

$$M_{ij} = \Phi_{ij} S_{ij} D_{ij} \quad (2)$$

Durch die Wahl der Multiplikation als Verknüpfungsoption wird gewährleistet, dass alle Bestandteile des Symmetriegrades in hohem Maße erfüllt sein müssen. Falls eine der Komponenten nahe an Null reicht, so wird auch der gesamte Symmetriegrad des entsprechenden Merkmals-Paares sehr klein. Bei einer Addition der Komponenten wäre dem nicht so: die zwei anderen Komponenten könnten den Verlust ausgleichen, was für das hier behandelte Problem nicht wünschenswert wäre.

2.2 Ergebnisse der Symmetrie-Detektion

Die Polarkoordinaten der Symmetrieachsen der potentiell symmetrischen Merkmals-Paare werden, gewichtet mit ihrem Symmetriegrad M , in einem zweidimensionalen Histogramm akkumuliert. Aus den Maxima dieses Histogramms lassen sich direkt die dominanten Symmetrien des betrachteten Bildes ableiten. Abbildung 3 zeigt dieses Histogramm für das Beispiel aus Abbildung 4. Die hier gezeigte Fassade ist ausschließlich durch vertikale Symmetrien gekennzeichnet. So zeigt das Histogramm ein globales Maximum bei $(\theta = 90^\circ, r = 391\text{pix})$ und mehreren lokale Nebenmaxima entlang der 90° Gitterlinie. Alle Spitzen, die durch mindestens drei Merkmals-Paare unterstützt werden, werden als Symmetrieachse für dieses Bild interpretiert.

Die Abbildung 4 f) zeigt alle fünf detektierten Symmetrieachsen im Überblick. Die Abbildungen a) – e) zeigen jeweils die Einzelergebnisse zusammen mit den konvexen Hüllen der für diese Symmetrie beteiligten Merkmalspunkte. Für dieses Beispiel wurden 1617 Merkmale detektiert und darunter 151 potentiell symmetrische Merkmals-Paare gefunden. Die Haupt-symmetrieachse (Abbildung 4 a) wird durch 34 Merkmals-Paare unterstützt.

Die detektierten Symmetrien konzentrieren sich auf den symmetrischen Teil der Fassade, andere Objekte des Bildes stören die Symmetrie-Detektion nicht. Die konvexe Hülle der an allen Symmetrien beteiligten Merkmale führt hier direkt zu der Bildregion, die durch symmetrische Strukturen gekennzeichnet ist.

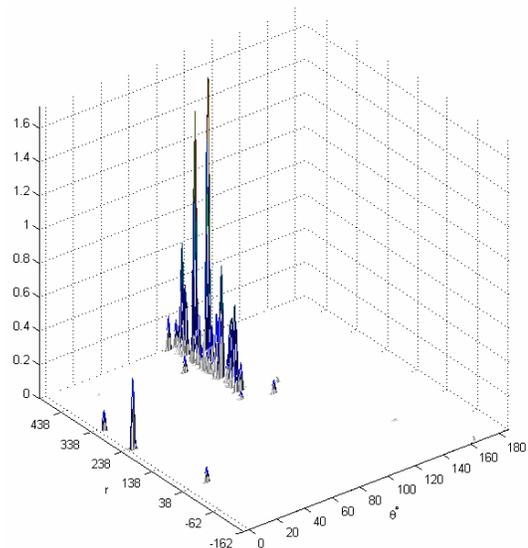


Abbildung 3: 2D – Histogramm der Polarkoordinatenrepräsentation der Symmetrieachsen für das Beispiel aus Abbildung 4.

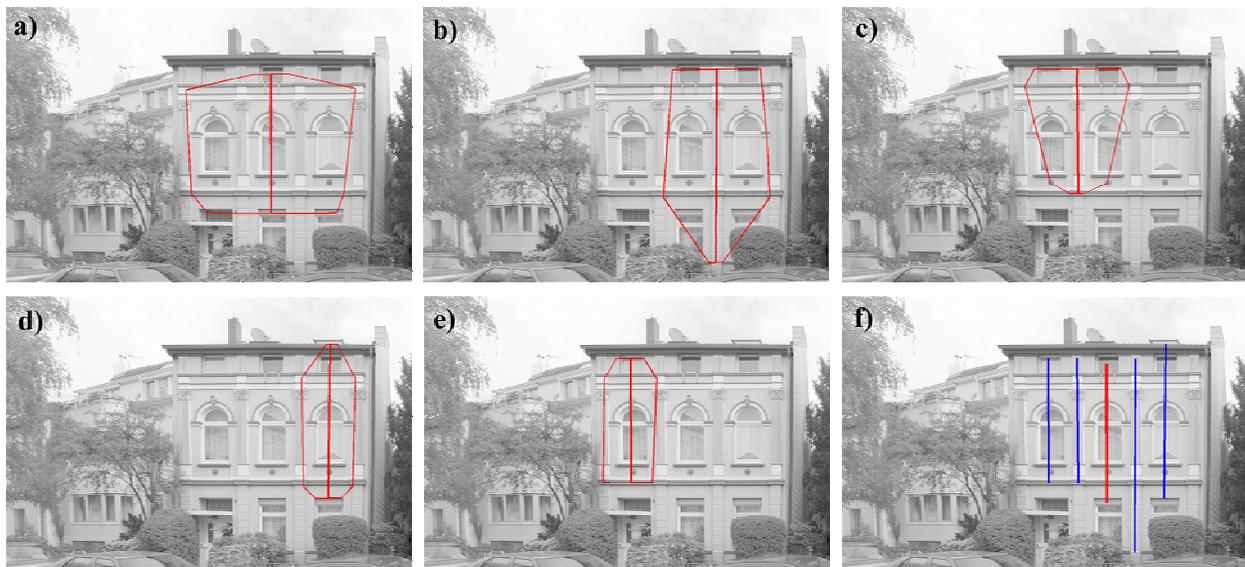


Abbildung 4: Ergebnisse der Symmetrie-Detektion. a) – e) Einzeldarstellung der gefundenen Symmetrien mit den konvexen Hüllen der beteiligten Merkmalspunkte. f) Gesamtdarstellung aller für dieses Beispiel gefundenen Symmetrien.

3 Detektion wiederholter Strukturen

Die von Loy und Eklundh vorgestellte Idee, Merkmale innerhalb eines Bildes einander zuzuordnen und die gefundenen Paare zu clustern, haben wir für die Detektion sich wiederholender Strukturen adaptiert. Dafür sind nur einige wenige Änderungen erforderlich.

Zunächst entfällt offensichtlich die Spiegelung der Deskriptoren. D. h. die im Bild detektierten Merkmalspunkte werden untereinander zugeordnet, um Paare potentiell gleicher Merkmale zu finden.

Die in 2.1 vorgestellten Gewichtungen werden angepasst. Das Winkel-Symmetrie-Gewicht Φ_{ij} vereinfacht sich damit zum Winkel-Gewicht $\Phi_{ij}^* \in [-1,1]$

$$\Phi_{ij}^* = \cos(\varphi_i - \varphi_j) \quad (3)$$

Womit diejenigen Merkmals-Paare den höchsten Support erhalten, die möglichst gleich orientiert sind. Die Gewichtungen in Abhängigkeit von Skalen-Differenz und Distanz bleiben unverändert. Damit ergibt sich die Bewertung der Qualität der Ähnlichkeit zweier Merkmale p_i und p_j wieder aus den Einzelgewichten als Repititionsgrad M^*

$$M_{ij}^* = \Phi_{ij}^* S_{ij} D_{ij} \quad (4)$$

Die Clustering der Merkmals-Paare über die Richtung ihre Verbindungslinie und den Betrag der Translation führt zu den dominanten Translationen im Bild und damit zu den wiederholten Strukturen.

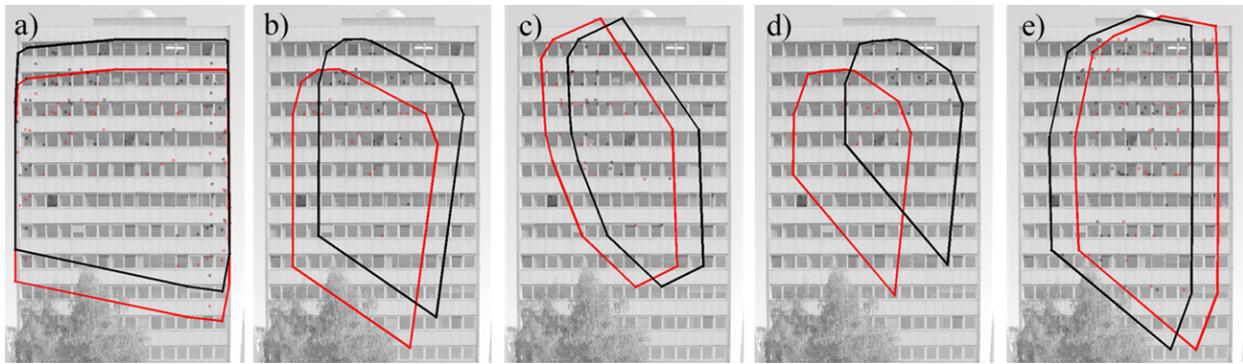


Abbildung 6: Ergebnisse der Detektion wiederholter Strukturen, in Form der ersten fünf detektierten wiederholten Gruppen. Dargestellt sind die beteiligten Merkmalspunkte und ihre konvexen Hüllen.

3.1 Ergebnisse der Detektion wiederholter Strukturen

Auch hier werden diejenigen Maxima des Histogramms als dominante Translation im Bild akzeptiert, die durch mindestens drei Merkmals-Paare unterstützt werden. Abbildung 6 zeigt die ersten fünf detektierten wiederholten Strukturen für dieses Beispiel. Dargestellt sind jeweils zwei Gruppen von Merkmalen und deren konvexe Hüllen. Alle Punkte einer Gruppe sind durch die gleiche Translation mit einem Punkt der anderen Gruppe verbunden.

Für dieses Beispiel wurden insgesamt 157 wiederholte Gruppen detektiert.² Zur Veranschaulichung dieser Ergebnisse zeigt Abbildung 5 alle detektierten Translationen als Plot von Translationsvektoren.³

Diese Darstellung zeigt deutlich die Regelmäßigkeit in den detektierten Translationen. Genau diese Regelmäßigkeit bzw. das Muster dahinter soll im Folgenden gefunden werden, um damit eine kompakte Beschreibung der Bildstruktur zu erhalten.

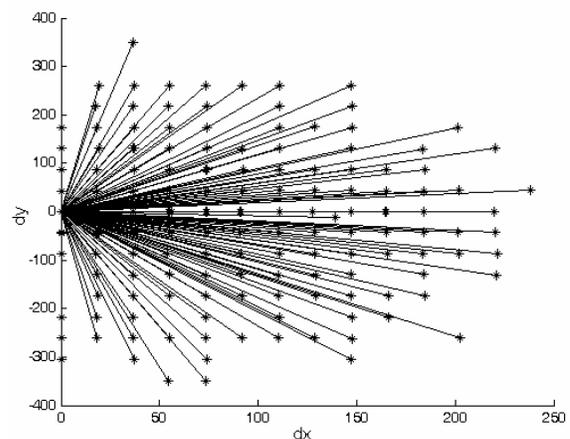


Abbildung 5: Die Translationsvektoren der 157 detektierten wiederholten Gruppen für das Bsp. aus Abbildung 6.

4 Ableitung einer kompakten Bildbeschreibung

Da wir mit entzerrten Bildern arbeiten, nehmen wir die Hauptrichtungen der Translationen als achsparallel an. Daher können wir die Suche nach einer geeigneten Basis getrennt auf die horizontale sowie vertikale Richtung beschränken.

Die beobachteten Translationen d_i ($i = 1 : n$) einer Richtung lassen sich somit als Linearkombinati-

² Wir haben unser Zuordnungskriterium mit $\text{distRatio} = 0.9$ sehr sensibel bzgl. Varianzen (Schatten, Gardinen etc.) der Fassadenobjekte gewählt. Damit führen auch noch relativ große Distanzen zwischen den Deskriptoren der Features zu einer positiven Zuordnung. Zu den Details der Parameter siehe [LOWE 2004].

³ Man beachte die unterschiedliche Skalierung der Achsen!

on von Basistranslationen v_k und der entsprechenden Koeffizienten α_k durch

$$d_i = \sum_{k=1}^K (\alpha_k \cdot v_k) + \varepsilon_i; \quad w_i \quad (5)$$

beschreiben. Für das synthetische Beispiel aus Abbildung 7 mit $K=2$, ließe sich die Fassade in horizontaler Richtung (c) durch

$$\begin{aligned} {}^c d_i &= {}^c \alpha_1 \cdot {}^c v_1 + {}^c \alpha_2 \cdot {}^c v_2 \\ \max |{}^c \alpha_1| &= 1 \quad \max |{}^c \alpha_2| = 2 \end{aligned} \quad (6)$$

beschreiben. Die Tiefe K der Basis-Hierarchie entspricht der Anzahl der Elemente der Linearkombination. A priori ist dieser Wert nicht bekannt, wir sind aber davon ausgegangen, dass typische städtische Fassaden in ihrer Komplexität den Wert 4 nicht überschreitet.

Weder die ganzzahligen Koeffizienten α_k noch die reelwertigen Basistranslationen v_k sind bekannt. Jede Beobachtung d_i ist mit einem Residuum ε_i und einem Gewicht w_i behaftet, womit bspw. Ausreißer kontrolliert werden können.

Gesucht ist eine Basis, bestehend aus K Basiselementen, durch die die im Bild beobachteten Translationen bestmöglich zu erklären sind, unter Minimierung der Residuen und der Komplexität K der Lösung.

Da für dieses Problem keine direkte Lösung von uns gefunden wurde, haben wir uns für ein heuristisches Verfahren entschieden. Dafür bestimmen wir zunächst die Differenzen zwischen allen beobachteten Translationen. Die Spitzen im Histogramm dieser 2. Differenzen der Positionen ergeben potentielle Kandidaten für die gesuchten Basistranslationen. Für diese n Kandidaten bilden wir alle

$M = \binom{n}{K}$ Kombinationen möglicher Basen v_K der Tiefe $K = 1 : 4$. Für jede dieser potentiellen Lö-

sungen ${}^j v_K$ ($j = 1 : M$) werden für jede Beobachtung d_i die Koeffizienten α_k bestimmt. Daraus ergibt sich für jede Lösung ${}^j v_K$ ein Residuen-Vektor ${}^j \varepsilon$.

Gesucht ist nun diejenige Lösung, die mit der geringsten Modellkomplexität die Beobachtungen am besten erklärt.

Aus der Informationstheorie lässt sich mit der minimalen Beschreibungslänge (MDL – Minimum Description Length) ein Kriterium ableiten, welches aussagt, wie viel Bit nötig werden, um eine bestimmte Information zu codieren. Wenn eine bestimmte Datenmenge sich durch ein kompaktes Modell beschreiben lässt, dann sind nur noch die Modellparameter und mögliche Abweichungen der Daten von diesem Modell zu codieren. Diese Überlegung führt zu dem in [Rissanen 1989] vorgeschlagenem MDL-Kriterium

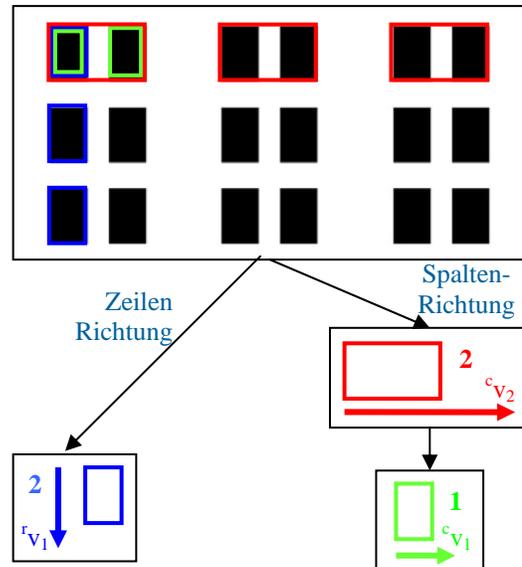


Abbildung 7: Synthetische Fassade, in vertikaler Richtung durch eine einfache Wiederholung gekennzeichnet. Die horizontale Struktur lässt sich durch eine einfache Translation der Elemente der Doppelfenster und deren Beziehung wiederum durch eine weitere zweifache Translation beschreiben. Die kompakte Bildbeschreibung besteht in horizontaler Richtung somit aus einer Hierarchie ($K=2$) von Basiselementen mit dem Betrag der Translation und der Anzahl der Wiederholungen.

$$MDL = -\log \prod_{i=1}^n P(x_i | \hat{\theta}) + \frac{K}{2} \log(n) \quad . \quad (7)$$

MDL gibt die zur Codierung des Modells $(\hat{\theta}, K)$ minimal benötigte Anzahl von Bits an. Gesucht ist dasjenige Modell $(\hat{\theta}, K)$, das mit der geringsten Komplexität K und der größten Datenwahrscheinlichkeit $\prod_{i=1}^n P(x_i | \hat{\theta})$ die beobachteten Daten x_i beschreibt.

Unter der Annahme normal verteilter Abweichungen der Beobachtungen vom Modell lässt sich das Kriterium darstellen mit

$$MDL = \frac{1}{2} \Omega + \frac{K}{2} \log(n) \quad . \quad (8)$$

Die Berücksichtigung von Ausreißern erfolgt nach Huber (siehe [FÖRSTNER 1989]) mit der Optimierungsfunktion

$$\rho(\varepsilon) = \begin{cases} T^2 & \text{falls } \left(\frac{\varepsilon}{\sigma}\right)^2 \geq T^2 \\ \left(\frac{\varepsilon}{\sigma}\right)^2 & \text{falls } \left(\frac{\varepsilon}{\sigma}\right)^2 < T^2 \end{cases} \quad \text{und} \quad \Omega = \sum_{i=1}^n \rho(\varepsilon_i) \quad . \quad (9)$$

Entsprechend den typischerweise für Hypothesentests verwendeten Signifikanzniveaus wählen wir den Schwellwert für Ausreißer zu $T = 3 \cdot \sigma$, mit $\sigma = 1.5$.

Die konvexen Hüllen der Merkmals-Paare, die das gewählte Modell v_K unterstützen, definieren die Bereiche im Bild die durch genau diese Basiselemente zu beschreiben sind. Damit sind wir, in Form von Basiselementen und den zugehörigen Bereichen im Bild, zu der von uns gesuchten mi-

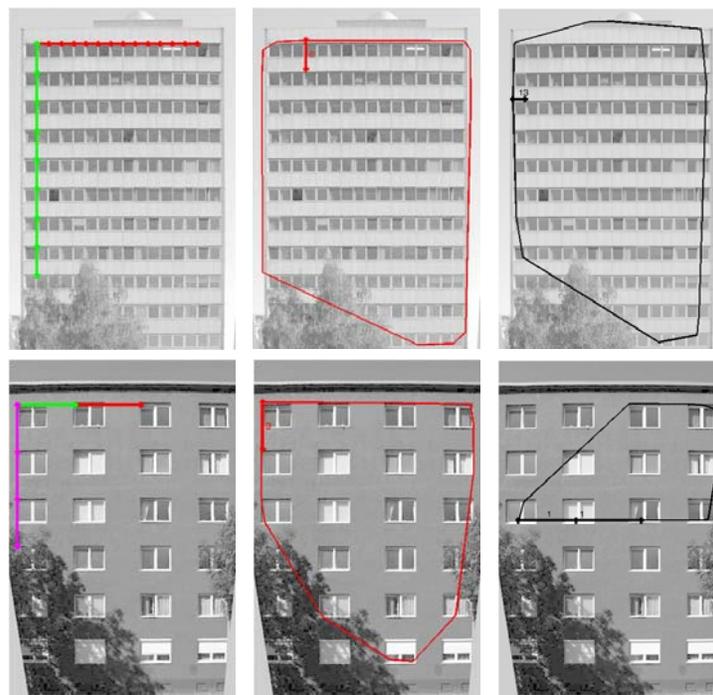


Abbildung 8: Ergebnisse der Bestimmung der kompakten Bildbeschreibung. Links: gefundenen Basiselemente, manuell platziert. Die Anzahl der Wiederholungen der Basisvektoren ist durch die bei der Bestimmung der Linearkombinationen der Basisvektoren für die Beobachtungen maximal auftretenden Koeffizienten gegeben. Mitte/ Rechts: Regionen die durch die vertikalen/ horizontalen Basiselemente zu beschreiben sind, mit dem jeweiligen Basisvektor.

nimalen Bildbeschreibung gelangt.

Abbildung 8 zeigt die Ergebnisse der Bestimmung der kompakten Bildbeschreibung für zwei Fassadenaufnahmen. Für das Beispiel aus Abbildung 6 wurden für beide Achsrichtungen einfache Basen ermittelt. Die konvexen Hüllen der an diesen Basen beteiligten Merkmale bedecken den gesamten Fassadenbereich, mit Ausnahme der vom Baum verdeckten Bereiche.

Abbildung 8 unten zeigt ein Beispiel einer hierarchischen Basis in horizontaler Richtung. Der Struktur der Fassade entsprechend, wurde ein Basiselement ermittelt, das die eng beieinander liegenden Fensterreihen verbindet und ein weiteres Basiselement, welches die Doppelfensterreihen verbindet.

5 Zusammenfassung und Ausblick

Wir haben gezeigt, wie der Ansatz von [LOY 2006] auf die Detektion multipler wiederholter Gruppen erweitert werden kann. Aus diesen von uns detektierten Translationen im Bild haben wir über ein Suchverfahren und das Kriterium der minimalen Beschreibungslänge ein Modell zur kompakten Bildbeschreibung abgeleitet.

Die gezeigten Beispiele beziehen sich auf Bilder, die fast ausschließlich den regelmäßigen Teil von Fassaden zeigen. Durch die sehr großzügige Wahl des Zuordnungskriteriums ist das Verfahren zum einen sehr sensibel für die wiederholten Objekte im regelmäßigen Teil der Fassade, aber zum anderen auch für ähnliche Merkmale am Rande derselben. Insofern ist die Bestimmung des Modells zur kompakten Bildbeschreibung noch weiter zu verfeinern, insbesondere im Hinblick auf die Robustheit gegen Störungen im Bild.

6 Literaturverzeichnis

- [FÖRSTNER 1989] FÖRSTNER, W., 1989: Image Analysis Techniques for Digital Photogrammetry. In: Vorträge der 42. Photogrammetrischen Woche an der Universität Stuttgart 1998, Schriftreihe Institut für Photogrammetrie der Universität Stuttgart, Heft 13, pp. 205-221.
- [LOWE 2004] LOWE, D. G., 2004: Distinctive Image Features from Scale-Invariant Keypoints. International Journal of Computer Vision, 60(2): 91–110.
- [LOY 2006] LOY, G. AND EKLUNDH, J.-O., 2006: Detecting Symmetry and Symmetric Constellations of Features. In Leonardis, A., Bischof, H., and Pinz, A., editors, Proceedings of ECCV 2006, volume II, pages 508–521. Springer-Verlag Berlin Heidelberg. LNCS 3952.
- [RISSANEN 1989] RISSANEN, J., 1989: Stochastic Complexity in Statistical Inquiry. Series in Computer Science - Vol. 15; World Scientific.
- [SCHAFFALITZKY 2000] SCHAFFALITZKY, F. AND ZISSERMAN, A., 2000: Planar Grouping for Automatic Detection of Vanishing Lines and Points. In: Image and Vision Computing, 2000, vol. 18, pages 647-658.
- [WENZEL 2006] WENZEL, S., 2006: Detektion wiederholter und symmetrischer Strukturen von Objekten in Bildern. Diplomarbeit. Institut für Geodäsie und Geoinformation, Fachbereich Photogrammetrie, Rheinische Friedrich-Wilhelms-Universität Bonn.