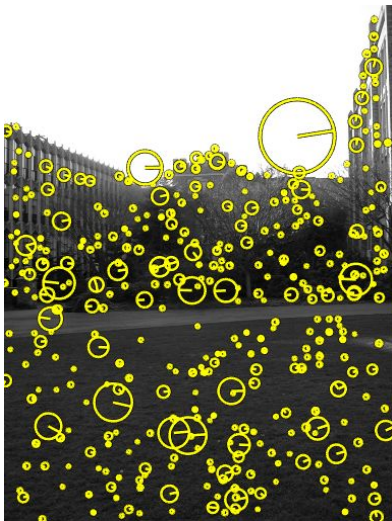


Aufgabe 4

A

- Show the output of `vl_plotframe` on the chosen image. What is the meaning of the size of the drawn circles and lines inside the circles?



<https://www.cs.ubc.ca/~lowe/papers/ijcv04.pdf>

Der Radius der Kreise bei jedem Feature Point gibt die Länge des Gradienten Vektors an welcher den Feature Punkt beschreibt. Der Zeiger im Kreis deutet die Richtung des Vektors an.

SIFT:

1) Erstellung des Scale Space/DoG

- 4 Oktaven zu je 5 scale Level, In jeder Oktave wurde das Bild im Vergleich zu vorherigen Oktave in jeder Dimension um die Hälfte reduziert, Jedes Bild innerhalb der Oktave ist mit einem je stärkeren gauss Filter gefaltet.
- Erstellung der Difference of Gaussian: (DoG); DoG -> sind jeweils immer die Differenzen der einzelnen gefilterten Bilder pro Oktave (so ähnlich wie laplace → Kantenfilter)

2) Nach Feature Points suchen

- Suche nach Maxima und Minima in DoG durch iteration aller Pixel und Nachbarschaftssuche, nicht nur im Image Space, sondern auch in der Nachbarschaft der oberen und unteren Skalierung Ebene in der derzeitigen Oktave, wird für alle Oktaven wiederholt.

3) Orientierung der Feature Points bestimmen

- Die Orientierung wird anhand von Gradienten in der Umgebung des Keypoints bestimmt. (Histogram mit jeweils 10° pro bin, bin mit der häufigsten anzahl ist der gradient für den KeyPoint)

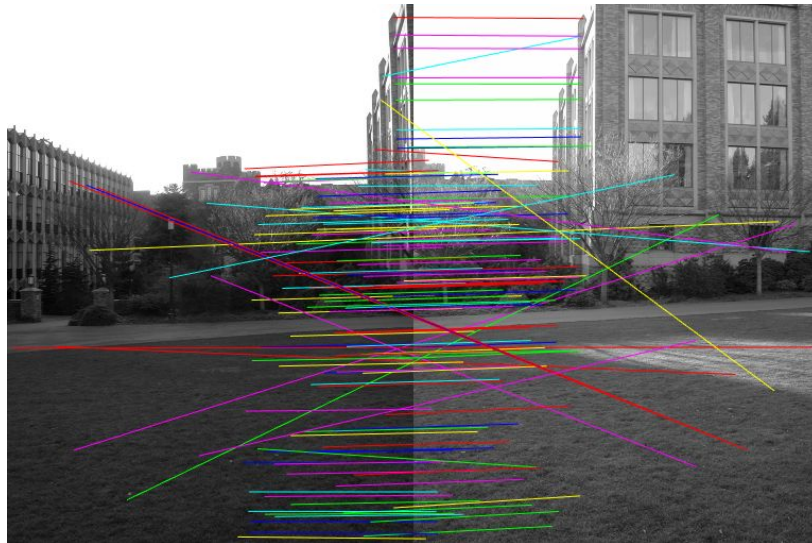
4) Den Fingerprint für den KeyPoint erstellen

- Um jeden Keypoint wird ein 16×16 Fenster betrachtet, welches in 4×4 Fenster mit je 16 Pixel unterteilt wird. Die Gradienten von diesen 16 Pixel werden in 16 Bins geteilt. die Bins sind in Kreisförmig angeordnet (360°) wobei jeder Bin 44° groß ist.



B

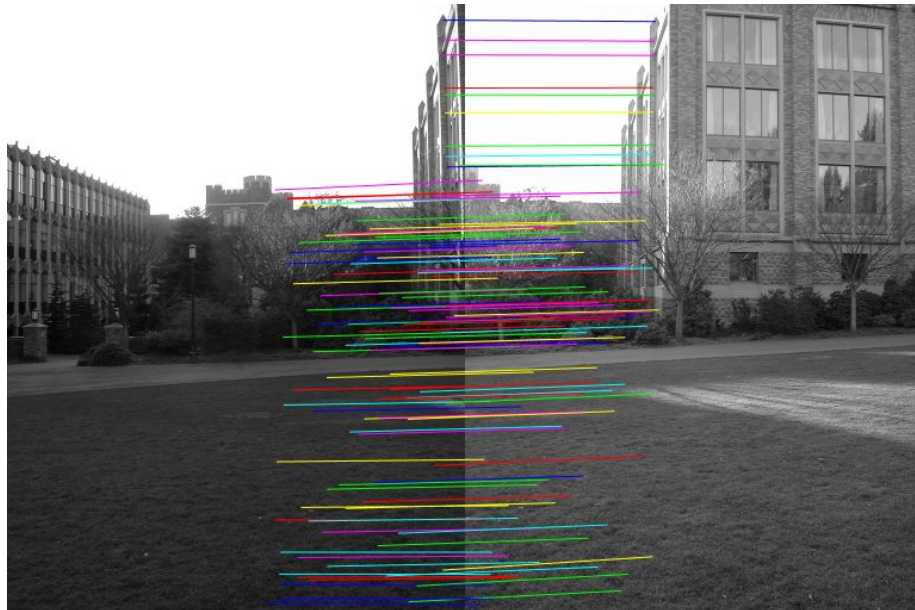
- Plot the matches after step 2. Describe in detail what `vl_ubcmatch` does.



Sucht gleiche KeyPoints im jeweils anderen Bild. Laut Lowe wird die geringste euklidische Distanz der invarianten description vectors hier verwendet. Ebenso wie laut der Dokumentation der Methode `vl_ubcmatch` zählen zwei Deskriptoren, wie zum Beispiel Description Vectors A und B, als Match wenn die euklidische Distanz, multipliziert mit einem gegebenen Threshold (default 1.5) nicht größer ist als A zu allen anderen Vektoren im anderen Bild.

5)

- Plot the matches of the inliers after step 4. What is the difference to the set of all putative matches we plotted before?

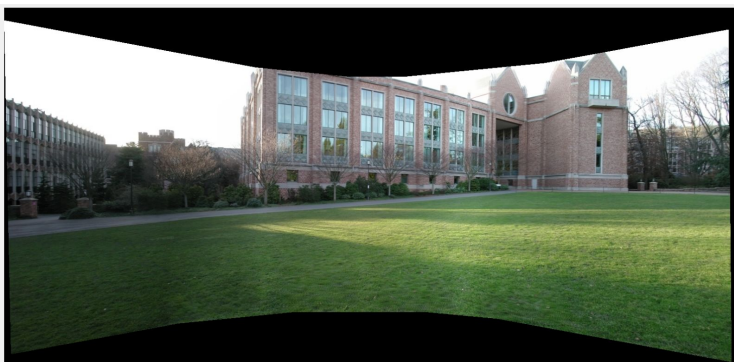


Nach RANSAC können die vermutlich falsche Matches (Outliers) entfernt werden. Statistisch gesehen werden Ausreißer aus der Gesamtmenge aller KeyPoint-Paare entfernt.

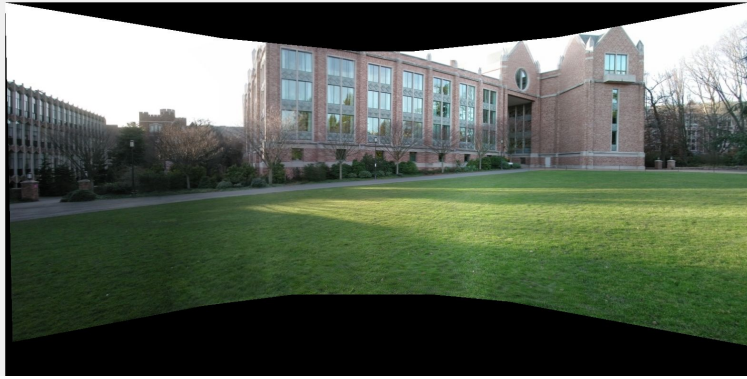
- Examine if the presented scheme of SIFT interest point detection and RANSAC-based alignment is invariant to changes in image rotation and scale. Thus, likewise to assignment 3, resize and rotate the second image and repeat the alignment procedure.

Scale Invariant	Ja, da Feature Points mit SIFT auf unterschiedlichen Ebenen gesucht werden
Rotations Invariant	Ja, durch die Berechnung der Orientierung des KeyPoints

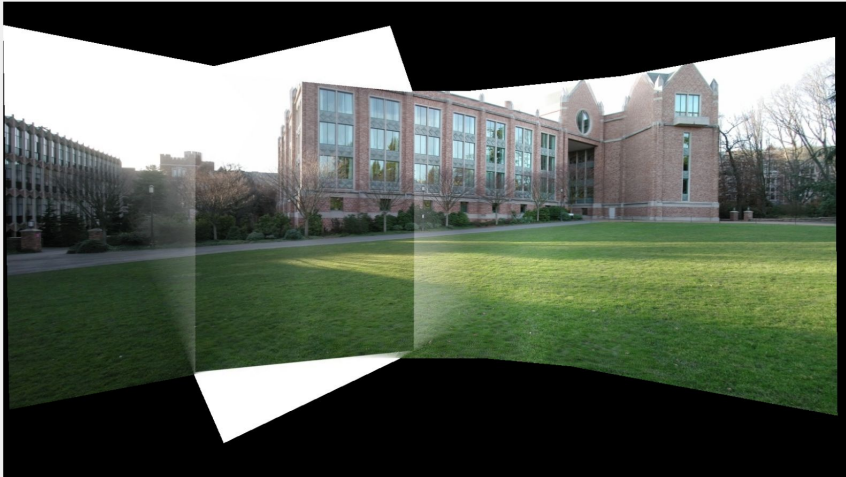
2. Bild ist größer:



2. Bild ist kleiner:



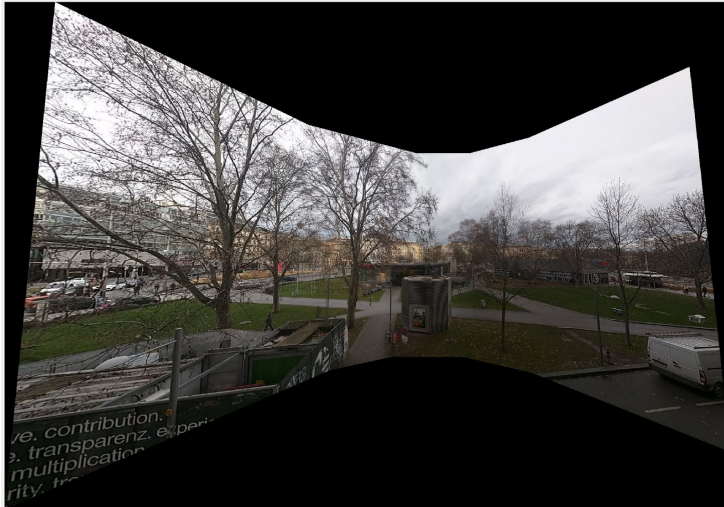
2. Bild rotiert: (Artefakte sind wegen dem weißem Rand zu sehen)



C

- Show and discuss the achieved results (with the two provided image sequences and your own sequence). The result might look quite realistic at a first glance but can you spot any errors by looking on details?

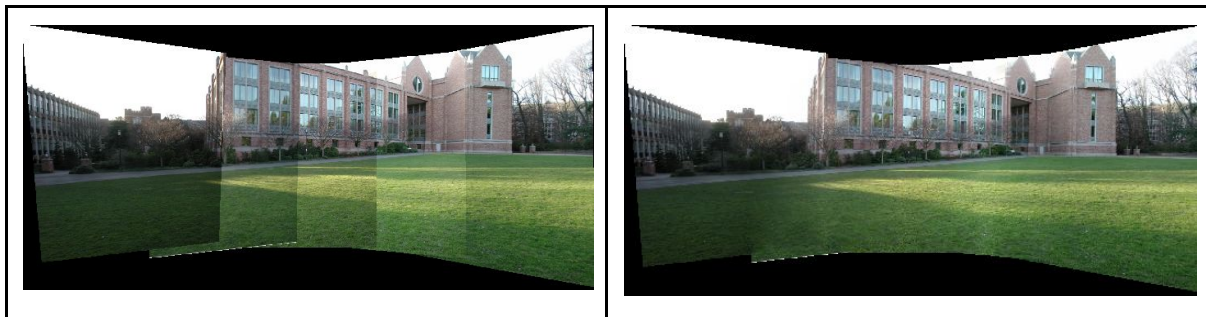




Durch unterschiedlich beleuchtete Bereiche in Szene sieht man nach dem feathering noch leicht die Kanten der einzelnen Bilder, da hier die Farbwerte von 2-3 Elnzelfotos abgebildet werden.

Auf der Wiese kann man blending Artefakte erkennen (hellere vertikale Pfade)
Auf dem eigenen Datensatz wurden die Bilder stärker verzerrt, vermutlich weil der Gesamtwinkel der Aufnahme größer ist

- Compare a result achieved with feathering to a result where no blending has been performed (i.e. the color values of only one image are taken for the stitched image). What is the difference of the two results?



- Kanten klar und deutlich sichtbar im nicht optimierten Bild
- Die unterschiedlichen Lichtverhältnissen sind mit feathering ausgeglichen.

Aufgabe 5

- In general, this approach should achieve a classification rate of around 60%. Show and discuss your results in the report. Show the confusion matrix and describe if there are classes which can be more easily identified than others. Additionally, describe if there are class pairs which are confused with one another more often than other pairs.

```

Command Window
current folder livingroom
current folder mountain
current folder office
current folder store
current folder street
Eval: 469 of 800 points have been classified correctly.
This is a positive rate of 58.625%.
fx K>>

```

Die Resultate des implementierten Algorithmus liegen konstant zwischen 50 und 60 Prozent. Diese Varianz kommt dadurch zustande, dass man beim Erstellen des Visual Vocabulary immer ein random sample aus 100 Punkten nimmt. Diese können manchmal akkurater sein und manchmal weniger.

evaluation									
8x8 double									
	1	2	3	4	5	6	7	8	9
1	30	0	11	31	4	9	9	6	
2	0	93	0	0	5	0	0	2	
3	7	0	45	17	1	19	10	1	
4	8	0	21	33	2	14	17	5	
5	1	10	0	0	75	1	3	10	
6	2	0	10	10	0	78	0	0	
7	8	2	5	16	3	4	53	9	
8	2	1	3	3	3	1	25	62	
9									

Die (oben abgebildete) confusion matrix zeigt auf der Diagonale die korrekt zugeordneten Klassen. Man sieht, dass die Klasse in Zeile/Spalte 2 die am öftesten korrekt zugeordnete Klasse ist (nämlich "forest"), gefolgt von Klasse 6 ("office") und 5 ("mountain").

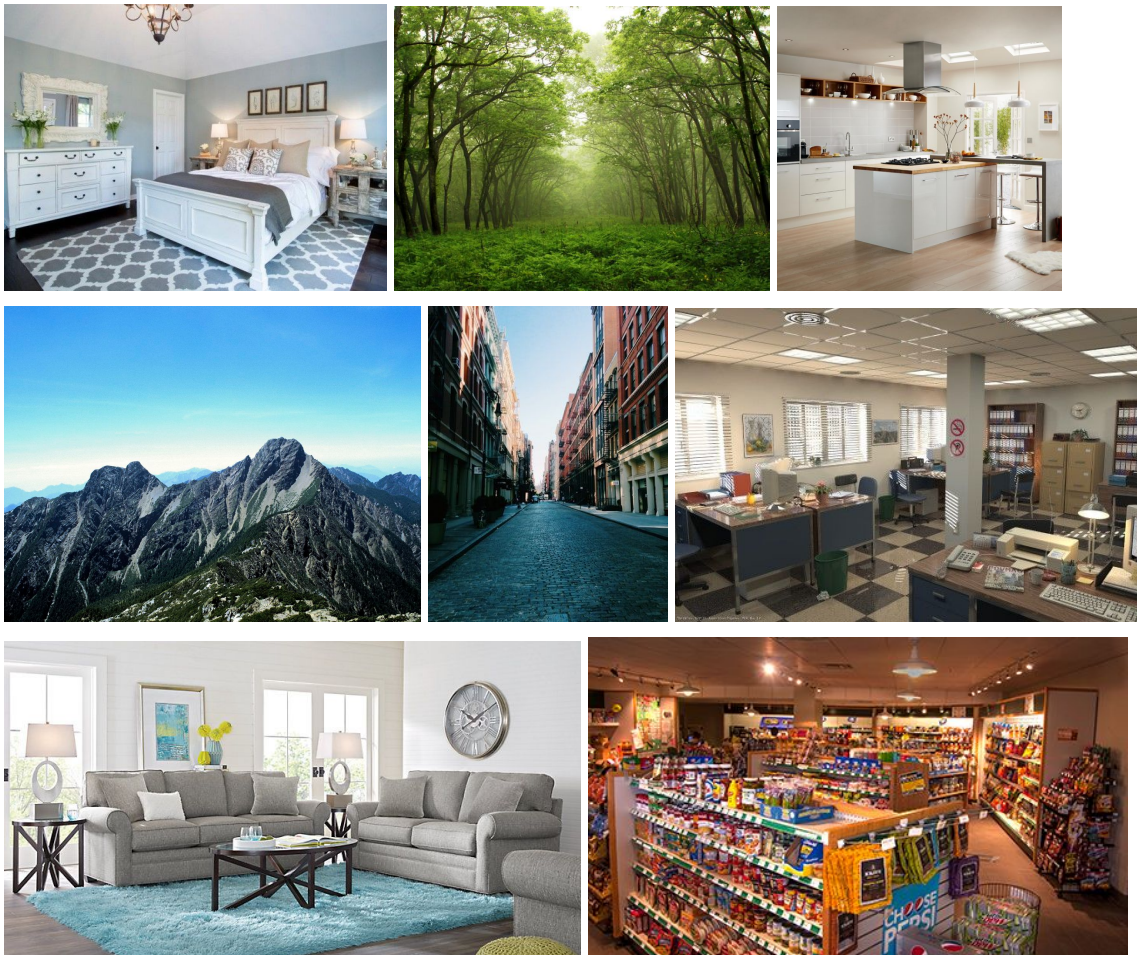
Was weiters bei den Ergebnissen heraussticht, ist, dass sehr viele (31!) eigentlich zur Klasse 1 ("bedroom") zugehörigen Bilder zur Klasse 4 ("living room") zugeordnet wurden. Dasselbe gilt für 25 Klasse 8 ("street") Bilder, die eigentlich zu Klasse 7 ("store") gehören sowie 21 Klasse 4 ("livingroom") Bilder, die inkorrekt zu Klasse 3 ("kitchen") zugeordnet wurden.

Aus diesen Ergebnissen kann man folgern, dass Naturszenen (forest, mountain) scheinbar für den Algorithmus leichter zu erkennen sind als Szenen mit menschengemachten Dingen darin. Das liegt womöglich daran, dass es in Naturszenen so viele unterschiedliche

Merkmale gibt. Dasselbe gilt für Office-Szenen, in diesen sind auch oft viele unterschiedliche Merkmale vorhanden.

Es ist auch weniger überraschend, dass die Innenszenen wie bedroom und living room sowie living room und kitchen vertauscht werden, da in Innenszenen eher gleiche Merkmale auftreten. Gerade Schlafzimmer und Wohnzimmer können sehr ähnlich aussehen und Wohnzimmer und Küche ineinander übergehen (vgl. Wohnküchen).

- Take some own test images and investigate if they can be correctly classified by your simple scene recognizer. For instance, make a photo of your own kitchen or living room and classify it. You should at least test one own photo from three of the eight categories.



Oben die verwendeten Testfotos der Reihe nach:
bedroom, forest, kitchen, mountain, street, office, living room, store

Im (besten) Test mit den eigenen Bildern, siehe Bilder unten, wurde wie erwartet forest und mountain korrekt klassifiziert. Interessanterweise wurde office in diesem Test nicht korrekt klassifiziert. Bedroom, store und street wurden ebenfalls korrekt erkannt.

