

Πολυτεχνική Σχολή

Τμήμα Μηχανικών Η/Υ & Πληροφορικής

**ΘΕΜΑΤΑ ΟΡΑΣΗΣ ΥΠΟΛΟΓΙΣΤΩΝ & ΓΡΑΦΙΚΗΣ**

**ΑΝΑΦΟΡΑ ΒΑΣΙΣΜΕΝΗ ΣΤΗΝ**

**3Η ΕΡΓΑΣΤΗΡΙΑΚΗ ΑΣΚΗΣΗ**

ΑΓΓΕΛΟΣ ΝΙΚΟΛΑΟΣ ΠΟΤΑΜΙΑΝΟΣ

Α.Μ. 1084537

up1084537@ac.upatras.gr

Πάτρα, 2024

ΑΣΚΗΣΗ 1.

1. **Ανίχνευση Ακροτάτων στον Χώρο Κλίμακας**Εδώ, δημιουργείται η πυραμίδα των εικόνων (Gaussian και DoG) και υπολογίζονται οι διαφορές των Γκαουσιανών.  
   Πιο συγκεκριμένα, από την αρχή του παραπάνω σχολίου μέχρι το σημείο πριν την επεξεργασία των ακροτάτων, ο κώδικας δημιουργεί τα Gaussian blurred images και τα DoG images. Αυτό περιλαμβάνει τις γραμμές:

%% Scale-Space Extrema Detection  
tic  
% original sigma and the number of actave can be modified. the larger  
% sigma0, the more quickly-smooth images  
sigma0=sqrt(2);  
octave=3;%6\*sigma\*k^(octave\*level)<=min(m,n)/(2^(octave-2))  
level=3;  
D=cell(1,octave);  
**for** i=1:octave  
D(i)=mat2cell(zeros(row\*2^(2-i)+2,colum\*2^(2-i)+2,level),row\*2^(2-i)+2,colum\*2^(2-i)+2,level);  
**end**  
% first image in first octave is created by interpolating the original one.  
temp\_img=kron(img,ones(2));  
temp\_img=padarray(temp\_img,[1,1],'replicate');  
figure(2)  
subplot(1,2,1);  
imshow(origin)  
%create the DoG pyramid.  
**for** i=1:octave  
…  
 **end**  
 D{i}=temp\_D;  
 temp\_img=temp\_img(1:2:**end**,1:2:**end**);  
 temp\_img=padarray(temp\_img,[1,1],'both','replicate');  
**end**  
toc

1. **Εύρεση Τοπικών Σημείων-Κλειδιών**

Μετά τον υπολογισμό των DoG, ο κώδικας εντοπίζει τα τοπικά ακρότατα στον χώρο κλίμακας. Σε αυτό το τμήμα γίνεται η σύγκριση κάθε pixel με τους 26 γείτονές του (9 στην προηγούμενη κλίμακα, 9 στην τρέχουσα και 9 στην επόμενη κλίμακα) για να εντοπιστούν τα τοπικά ακρότατα. Στη συνέχεια εφαρμόζονται έλεγχοι για την απόρριψη σημείων που δεν έχουν επαρκή αντίθεση ή βρίσκονται σε άκρα εικόνων.

%% Keypoint Localisation  
% search each pixel in the DoG map to find the extreme point  
tic  
interval=level-1;  
number=0;  
**for** i=2:octave+1  
 number=number+(2^(i-octave)\*colum)\*(2\*row)\*interval;  
**end**  
extrema=zeros(1,4\*number);  
flag=1;  
**for** i=1:octave  
…  
…  
…  
%% accurate keypoint localization   
%eliminate the point with low contrast or poorly localised on an edge  
% x:|,y:-- x is for vertial and y is for horizontal  
% value comes from the paper.  
tic  
threshold=0.1;  
r=10;  
extr\_volume=length(extrema)/4;  
[m,n]=size(img);  
secondorder\_x=conv2([-1,1;-1,1],[-1,1;-1,1]);  
secondorder\_y=conv2([-1,-1;1,1],[-1,-1;1,1]);  
**for** i=1:octave  
 **for** j=1:level  
 test=D{i}(:,:,j);  
 temp=-1./conv2(test,secondorder\_y,'same').\*conv2(test,[-1,-1;1,1],'same');  
 D{i}(:,:,j)=temp.\*conv2(test',[-1,-1;1,1],'same')\*0.5+test;  
 **end**  
**end**  
**for** i=1:extr\_volume  
…  
…  
…  
subplot(2,2,4);  
imshow(origin)  
hold on  
plot(ry,rx,'b+');  
toc

Εδώ γίνεται η λεπτομερής εύρεση ακριβών θέσεων των σημείων κλειδιών και απόρριψη των σημείων που δεν πληρούν τα κριτήρια αντίθεσης και καμπυλότητας της επιφάνειας.  
  
**3. Διόρθωση Κατεύθυνσης**  
Μετά τον εντοπισμό των σημείων κλειδιών, στο επόμενο στάδιο αποδίδεται προσανατολισμός σε κάθε σημείο-κλειδί, ώστε ο περιγραφέας που θα υπολογιστεί αργότερα να είναι αμετάβλητος ως προς την περιστροφή.

%% Orientation Assignment(Multiple orientations assignment)  
tic  
kpori=zeros(1,36\*extr\_volume);  
minor=zeros(1,36\*extr\_volume);  
f=1;  
flag=1;  
**for** i=1:extr\_volume  
…  
…  
…  
 **end**  
**end**  
idx= minor==0;  
minor(idx)=[];  
extrema=minor;  
% delete unsearchable points and add minor orientation points  
idx= kpori==0;  
kpori(idx)=[];  
extr\_volume=length(extrema)/4;  
toc

1. **Περιγραφέας Σημείων-Κλειδιών**

Αφού πλέον κάθε σημείο κλειδί έχει θέση, κλίμακα και προσανατολισμό, ακολουθεί ο υπολογισμός του περιγραφέα. Ο περιγραφέας SIFT βασίζεται στην κατασκευή προσανατολισμένων ιστογραμμάτων κλίσης σε υποπεριοχές γύρω από το σημείο-κλειδί. Εδώ υπολογίζεται το 128-διάστατο διάνυσμα που περιγράφει τοπικά το σημείο κλειδί. Με αυτό τον τρόπο τα σημεία μπορούν να αντιστοιχιστούν μεταξύ διαφορετικών εικόνων.

%% keypoint descriptor  
tic  
d=4;% In David G. Lowe experiment,divide the area into 4\*4.  
pixel=4;  
feature=zeros(d\*d\*8,extr\_volume);  
**for** i=1:extr\_volume  
 descriptor=zeros(1,d\*d\*8);% feature dimension is 128=4\*4\*8;  
 width=d\*pixel;  
 %x,y centeral point and prepare for location rotation  
 x=floor((extrema(4\*(i-1)+3)-1)/(n/(2^(extrema(4\*(i-1)+1)-2))))+1;  
 y=mod((extrema(4\*(i-1)+3)-1),m/(2^(extrema(4\*(i-1)+1)-2)))+1;  
 z=extrema(4\*(i-1)+2);  
 **if**((m/2^(extrema(4\*(i-1)+1)-2)-pixel\*d\*sqrt(2)/2)>x&&x>(pixel\*d/2\*sqrt(2))&&(n/2^(extrema(4\*(i-1)+1)-2)-pixel\*d/2\*sqrt(2))>y&&y>(pixel\*d/2\*sqrt(2)))  
 sub\_x=(x-d\*pixel/2+1):(x+d\*pixel/2);  
 sub\_y=(y-d\*pixel/2+1):(y+d\*pixel/2);  
 sub=zeros(2,length(sub\_x)\*length(sub\_y));  
 j=1;  
 **for** p=1:length(sub\_x)  
 **for** q=1:length(sub\_y)  
 sub(:,j)=[sub\_x(p)-x;sub\_y(q)-y];  
 j=j+1;  
 **end**  
 **end**  
  
…  
…  
…  
index=find(sum(feature));  
feature=feature(:,index);  
toc

ΑΣΚΗΣΗ 2.

**1. Ανίχνευση Ακροτάτων στον Χώρο Κλίμακας (Scale-Space Extrema Detection)**

**Βήματα:**

1. **Δημιουργία του Χώρου Κλίμακας:**
   * Για την αρχική εικόνα I(x, y), δημιουργούμε πολλαπλές "οκτάβες" (octaves).
   * Κάθε οκτάβα περιλαμβάνει σειρά εικόνων λείων με διαφορετικές τιμές σ (διάφορα επίπεδα κλίμακας).
   * Οι λείες εικόνες προκύπτουν με συνέλιξη της αρχικής εικόνας με Γκαουσιανά φίλτρα διαφορετικής διασποράς σ.
2. **Υπολογισμός Διαφοράς Γκαουσιανών (DoG):**
   * Για κάθε οκτάβα, από τις λείες εικόνες αφαιρείται η μία από την επόμενη, ώστε να προκύψουν οι εικόνες Difference-of-Gaussian (DoG).
   * Κάθε DoG εικόνα δημιουργείται χρησιμοποιώντας ένα σταθερό παράγοντα κλίμακας k.
3. **Προετοιμασία για Εντοπισμό Τοπικών Ακροτάτων:**
   * Οι DoG εικόνες οργανώνονται σε πυραμίδα με οκτάβες και επίπεδα.
   * Αυτό θα χρησιμοποιηθεί στο επόμενο στάδιο για ανίχνευση ακροτάτων.

**2. Εύρεση Τοπικών Σημείων-Κλειδιών (Keypoint Localization)**

**Βήματα:**

1. **Εντοπισμός Τοπικών Ακροτάτων:**
   * Κάθε εικονοστοιχείο σε κάθε DoG εικόνα συγκρίνεται με τους 8 γείτονές του στην ίδια κλίμακα και τους 18 γείτονες στις γειτονικές κλίμακες (συνολικά 26 γείτονες).
   * Ένα εικονοστοιχείο θεωρείται ακρότατο (τοπικό μέγιστο ή ελάχιστο) αν η τιμή του είναι είτε μεγαλύτερη είτε μικρότερη από όλες αυτές των 26 γειτόνων.
2. **Ακριβής Προσδιορισμός Θέσης Σημείου Κλειδιού:**
   * Χρησιμοποιείται μία προσέγγιση με επέκταση Taylor γύρω από την αρχική θέση ακροτάτου, ώστε να προσδιοριστεί η ακριβής θέση του στο συνεχές.
   * Με αυτόν τον τρόπο βρίσκουμε πιο ακριβή θέση σε υποεικονοστοιχειακό επίπεδο.
3. **Απόρριψη Ακατάλληλων Σημείων:**
   * Απορρίπτονται σημεία χαμηλής αντίθεσης (threshold στο μέγεθος του DoG).
   * Απορρίπτονται σημεία που βρίσκονται πάνω σε άκρα (edges), χρησιμοποιώντας τον λόγο των ιδιοτιμών του Hessian (Edge response test).

**3. Διόρθωση Κατεύθυνσης (Orientation Assignment)**

**Βήματα:**

1. **Υπολογισμός Τοπικών Κλίσεων γύρω από το Σημείο Κλειδί:**
   * Υπολογίζεται η κλίση μεγέθους m(x, y) και κατεύθυνσης θ(x, y) σε μια γειτονιά γύρω από το σημείο κλειδί, στην κατάλληλη κλίμακα.
2. **Κατασκευή Ιστογράμματος Κατευθύνσεων:**
   * Τα διανύσματα κλίσης τακτοποιούνται σε ένα ιστογράφημα 36 bins που καλύπτουν όλες τις πιθανές γωνίες (0° έως 360°).
   * Κάθε μέγεθος κλίσης σταθμίζεται και από έναν Γκαουσιανό πυρήνα ανάλογα με την απόσταση από το σημείο κλειδί.
3. **Εύρεση Κύριας και Συμπληρωματικών Κατευθύνσεων:**
   * Η κύρια κατεύθυνση ορίζεται ως η γωνία που αντιστοιχεί στο μέγιστο του ιστογράμματος.
   * Όποιες άλλες κορυφές στο ιστογράφημα που ξεπερνούν το 80% της κύριας κορυφής οδηγούν στη δημιουργία επιπλέον σημείων κλειδιού με τις αντίστοιχες κατευθύνσεις.

**4. Υπολογισμός του Τοπικού Περιγραφέα Σημείου Κλειδιού (Keypoint Descriptor)**

**Βήματα:**

1. **Κανονικοποίηση Έναντι Κλίμακας και Προσανατολισμού:**
   * Η περιοχή γύρω από το σημείο κλειδί περιστρέφεται σύμφωνα με την κύρια κατεύθυνση ώστε ο περιγραφέας να είναι αμετάβλητος ως προς την περιστροφή.
   * Λαμβάνεται η κατάλληλη κλίμακα γύρω από το σημείο ώστε να είναι αμετάβλητος και ως προς την αλλαγή κλίμακας.
2. **Δημιουργία Υποπεριοχών και Υπολογισμός Ιστογραμμάτων Κλίσης:**
   * Χωρίζεται η περιοχή γύρω από το σημείο κλειδί σε ένα πλέγμα 4x4 υποπεριοχών.
   * Για κάθε υποπεριοχή, υπολογίζεται ένα ιστογράφημα 8 προσανατολισμών.
   * Συνολικά προκύπτουν 4x4x8 = 128 τιμές, που αποτελούν το χαρακτηριστικό διάνυσμα του περιγραφέα.
3. **Κανονικοποίηση του Περιγραφέα:**
   * Το 128-διάστατο διάνυσμα κανονικοποιείται ώστε να είναι ανεξάρτητο από την παγκόσμια φωτεινότητα ή την αντίθεση.
   * Αποκόπτονται μεγάλες τιμές (clipping), ακολουθεί εκ νέου κανονικοποίηση για επιπλέον σταθερότητα.

ΑΣΚΗΣΗ 3  
  
Σύμφωνα με τον κώδικα SIFT\_feature.m που δόθηκε και την τυπική υλοποίηση του αλγορίθμου SIFT, ο αριθμός των επιπέδων (levels) ανά οκτάβα είναι συνήθως 3. Στο συγκεκριμένο παράδειγμα κώδικα που εξετάσαμε, χρησιμοποιείται η εντολή:

**level** = 3;

η οποία ορίζει ότι ανά οκτάβα παράγονται 3 επίπεδα κλίμακας για τον υπολογισμό των διαφορών Gaussians (DoG). Έτσι, η απάντηση είναι ότι χρησιμοποιούνται 3 επίπεδα ανά οκτάβα.  
  
**ΑΣΚΗΣΗ 4**  
  
**1. Παρεμβολή με Πολυωνική Προσέγγιση (Polynomial Fitting)**

Μετά τον εντοπισμό ενός σημείου κλειδιού σε συγκεκριμένη κλίμακα, μπορούμε να εφαρμόσουμε μία πολυωνική προσαρμογή (π.χ. τριών διαστάσεων για x, y, κλίμακα σ) των τιμών της DoG. Με αυτόν τον τρόπο, προσεγγίζουμε τη συνάρτηση απόκρισης γύρω από το σημείο και επιλύοντας τις μερικές παραγώγους (διαφορές) βρίσκουμε τη θέση του μέγιστου ή ελάχιστου με υποεικονοστοιχειακή (sub-pixel) και υποκλίμακα (sub-scale) ακρίβεια.

**Πλεονεκτήματα:**

* Αυξημένη ακρίβεια στον προσδιορισμό της θέσης και της κλίμακας.
* Μειώνει την επίδραση του θορύβου, καθώς η πολυωνική προσαρμογή λειαίνει τις τιμές.

**Μειονεκτήματα:**

* Αυξημένο υπολογιστικό κόστος σε σύγκριση με την απλή διακριτοποίηση.
* Ευαισθησία στην επιλογή του πολυωνικού βαθμού: πολύ χαμηλός βαθμός δεν προσφέρει σημαντική βελτίωση, πολύ υψηλός μπορεί να οδηγήσει σε υπερπροσαρμογή (overfitting).

**2. Χρήση Μεθόδου Επέκτασης Taylor (Taylor Expansion)**  
Στην υλοποίηση του SIFT χρησιμοποιείται ήδη μία τρισδιάστατη επέκταση Taylor γύρω από το αρχικά εντοπισμένο ακρότατο. Μπορούμε να εφαρμόσουμε μια πιο ακριβή ή επαναληπτική διαδικασία προσέγγισης Taylor, επεκτείνοντας τους υπολογισμούς ή εκτελώντας περισσότερα βήματα βελτιστοποίησης.

**Πλεονεκτήματα:**

* Εφαρμόζεται σχετικά εύκολα καθώς ήδη χρησιμοποιείται στο βασικό SIFT.
* Δίνει κλειστή μορφή λύσεων για τη μετατόπιση κλίμακας και θέσης.

**Μειονεκτήματα:**

* Εξαρτάται από την ορθότητα του μοντέλου Taylor· αν η συνάρτηση DoG δεν είναι καλά προσεγγίσιμη από ένα πολυώνυμο χαμηλού βαθμού, η ακρίβεια περιορίζεται.
* Προσθέτει επιπλέον επαναλήψεις και υπολογιστικό κόστος.

**3. Χρήση Μεθόδων Βελτιστοποίησης (Optimization Techniques)**  
Εφαρμογή αριθμητικών μεθόδων βελτιστοποίησης, όπως ο αλγόριθμος Newton-Raphson ή Levenberg-Marquardt, πάνω στην συνάρτηση DoG στον χώρο κλίμακας. Ξεκινούμε από την αρχική εκτίμηση και βελτιώνουμε επαναληπτικά τη θέση και την κλίμακα έως ότου συγκλίνουμε σε ένα τοπικό μέγιστο/ελάχιστο.

**Πλεονεκτήματα:**

* Μεγάλη ευελιξία: μπορούμε να συνδυάσουμε διαφορετικές συναρτήσεις κόστους, τακτικοποιήσεις ή περιορισμούς.
* Δυνατότητα εύρεσης πιο ακριβούς λύσης εφόσον το πρόβλημα είναι καλά ορισμένο.

**Μειονεκτήματα:**

* Αυξημένο υπολογιστικό κόστος.
* Πιθανότητα παγίδευσης σε τοπικά ακρότατα εάν η αρχική εκτίμηση ή το μοντέλο δεν είναι επαρκώς καλές.
* Απαιτεί καλή ρύθμιση παραμέτρων (π.χ. μήκους βήματος, κριτήρια σύγκλισης).

**4. Αυξημένη Πυκνότητα Δειγματοληψίας στην Κλίμακα (Over-Sampling in Scale Space)**  
Αντί να χρησιμοποιείται ένας συγκεκριμένος παράγοντας k για την αλλαγή κλίμακας μεταξύ των επιπέδων, μπορούμε να παράγουμε περισσότερες ενδιάμεσες εικόνες (π.χ. 4 ή 5 επίπεδα ανά οκτάβα αντί για 3) ώστε να έχουμε πιο πυκνή δειγματοληψία και άρα μικρότερο κενό μεταξύ διαδοχικών κλιμάκων.

**Πλεονεκτήματα:**

* Απλή μέθοδος που δεν απαιτεί περίπλοκη μαθηματική μοντελοποίηση.
* Επιτυγχάνει καλύτερη κατά προσέγγιση ακρίβεια χωρίς πολύπλοκη βελτιστοποίηση.

**Μειονεκτήματα:**

* Αυξάνει σημαντικά τον υπολογισμό (περισσότερες εικόνες, περισσότερος χρόνος επεξεργασίας).
* Δεν δίνει πραγματικά υποκλίμακα ακρίβεια, απλώς πυκνώνει το "πλέγμα" στον άξονα κλίμακας.

**5. Προσέγγιση με Πολυδιάστατη Ενδοδιάμεση Παρεμβολή (Multi-Dimensional Interpolation)**  
Χρήση ενδοδιάμεσης παρεμβολής (interpolation) με ανώτερες spline συναρτήσεις στον τρισδιάστατο χώρο (x, y, κλίμακα). Με αυτό τον τρόπο μπορούμε να εκτιμήσουμε τις τιμές DoG σε μη ακέραια επίπεδα κλίμακας και να βρούμε ακριβέστερα το ακρότατο.

**Πλεονεκτήματα:**

* Πολύ ομαλή προσέγγιση της συνάρτησης.
* Δυνατότητα πολύ υψηλής ακρίβειας.

**Μειονεκτήματα:**

* Πολύπλοκη υλοποίηση και ακόμη υψηλότερο υπολογιστικό κόστος.
* Ευαισθησία σε εξωγενείς παραμορφώσεις (π.χ. θόρυβος, κακή αρχικοποίηση).

**ΑΣΚΗΣΗ 5**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| liftingbody | moon | tire | rice | westconcordorthophoto |
|  |  |  |  |  |

Επιλέγω ασπρόμαυρες φωτογραφίες λόγω περιορισμού από το directory: \Matlab\toolbox\images\imdata :  
liftingbody.png,moon.tif, tire.tif, rice.png, westconcordorthophoto.png

1. **liftingbody.png:**
2. **moon.tif:**Στην εικόνα του φεγγαριού, τα keypoints εντοπίζονται κυρίως κατά μήκος του τόξου που διαχωρίζει το φωτεινό τμήμα της σελήνης από το σκοτεινό φόντο.

Επίσης παρατηρείται εντοπισμός σημείων σε περιοχές με κρατήρες ή επιφανειακές λεπτομέρειες, όπου τοπικές μεταβολές του φωτισμού είναι έντονες.

Ο αλγόριθμος SIFT βρίσκει λιγότερα keypoints σε μεγάλες ομαλές περιοχές (π.χ. η σκιά/σκοτεινό φόντο), καθώς δεν εμφανίζονται μεγάλες μεταβολές έντασης.

Οι διαφορές στο πλήθος και την κατανομή (πράσινο/μπλε) υποδηλώνουν είτε διαφορές στις ρυθμίσεις είτε την ευαισθησία του αλγορίθμου στην αντίθεση. Σε υψηλότερο κατώφλι αντίθεσης, θα εμφανιστούν λιγότερα σημεία, εστιάζοντας στα πιο εμφανή χαρακτηριστικά (κρατήρες/όριο φωτεινότητας).

1. **tire.tif:**Το SIFT εντοπίζει πολλά σημεία γύρω από τα όρια της ζάντας, τις εσοχές, τις βίδες και τις αλλαγές φωτεινότητας στο ελαστικό.

Οι κυκλικές συμμετρίες και οι έντονες τοπικές μεταβολές οδηγούν στην ανίχνευση πολλών keypoints, ιδιαίτερα σε περιοχές με μεταλλική αντανάκλαση.  
Η διαφορά μεταξύ πράσινου/μπλε συνόλου καταδεικνύει ότι μικρές αλλαγές στις παραμέτρους ενδέχεται να προκρίνουν διαφορετικό πλήθος και ακριβή τοποθεσία σημείων, ωστόσο το γενικό μοτίβο παραμένει παρόμοιο (συγκέντρωση σημείων σε σημαντικές μεταβολές).

1. **rice.png**Τα keypoints διασπείρονται σε περιοχές όπου η αντίθεση μεταξύ των κόκκων και του φόντου είναι έντονη.

Λόγω της ομοιομορφίας του σχήματος (πολλοί κόκκοι με παρόμοια εμφάνιση), εμφανίζονται πάρα πολλά σημεία, πιθανόν σε άκρα των κόκκων.

Η ταύτιση πολλών παρόμοιων χαρακτηριστικών είναι αναμενόμενη, αλλά αυτό μπορεί να δυσκολέψει τη μοναδικότητα του περιγραφέα.

Οι διαφορές πράσινου/μπλε συνόλου σημείων πιθανόν οφείλονται στον τρόπο που φιλτράρεται ο θόρυβος ή στη ρύθμιση της παραμέτρου αντίθεσης. Σε περιβάλλον γεμάτο επαναλαμβανόμενα μοτίβα, μια πιο αυστηρή ρύθμιση μειώνει τον αριθμό των keypoints.

1. **westconcordorthophoto.png**