

Άσκηση 1 στο Μάθημα Ρομποτικά Συστήματα

Θέμα: Ρομποτική Οδομετρία

Τμήμα: Η.Μ. & Τ.Υ. Τομέας Σ.Α.Ε.

Ραπτάκης Αθανάσιος - Α.Μ: 6623 Κακούρος Νίκος - Α.Μ:6519

Σκοπός της άσκησης:

Σκοπός της άσκησης είναι να εκτιμήσουμε την τροχιά ενός κινούμενου ρομπότ βασισμένοι στην οπτική πληροφορία που αποκτά από οπτικούς αισθητήρες-κάμερες που βρίσκονται προσαρμοσμένοι πάνω του.

Η διαδικασία της εκτίμησης της τροχιάς που θα ακολουθήσουμε βασίζεται στην παραδοχή ότι οι εικόνες που καταγράφουν οι κάμερες του ρομπότ αλληλοεπικαλύπτονται. Αυτό το γεγονός μας επιτρέπει να ανιχνεύσουμε τα κοινά σημεία ανάμεσα σε δύο διαδοχικές λείψεις χρησιμοποιώντας τον αλγόριθμο SIFT. Έπειτα θα εκτιμήσουμε την σχετική μετατόπιση της κάμερας στο χώρο με τη χρήση affine γεωμετρικών μετασχηματισμών.

- Ο αλγόριθμος που υλοποιήθηκε είναι:

1. Ανίχνευση feature points μέσω του αλγορίθμου SIFT στις εικόνες I_N και I_{N+1}
2. Υπολογισμός του affine γεωμετρικού μετασχηματισμού και ανάστροφου γεωμετρικού μετασχηματισμού μεταξύ των δύο διαδοχικών εικόνων.
3. Εξάγουμε την μετατόπιση της κάμερας T_x , T_y από τον πίνακα γεωμετρικού μετασχηματισμού που υπολογίσαμε.
4. Χρησιμοποιώντας τον ανάστροφο γεωμετρικό μετασχηματισμό ενώνουμε τις εικόνες I_N και I_{N+1} σχηματίζοντας το πανόραμα των εικόνων.
5. Η τροχιά του ρομπότ υπολογίζεται συσσωρεύοντας τη διαδοχική μετατόπιση της κάμερας T_x , T_y η οποία υποθέτουμε απριόρι ότι βρίσκεται στο κέντρο της εικόνας I_{N+1} .
6. Κατά την εκτέλεση του αλγορίθμου λαμβάνουμε ως δεδομένο ότι το ρομπότ κινείται μόνο σε ένα επίπεδο.

- Στο τέλος του αλγορίθμου παράγονται:

1. το **διάγραμμα της τροχιάς** στο (x,y) επίπεδο.
2. Το διάγραμμα του προσανατολισμού του κινούμενου οχήματος.
3. Τα διαγράμματα γραμμικής και γωνιακής ταχύτητας.
4. Παράγεται βίντεο με την διαδικασία δημιουργίας του πανοράματος.
5. Παράγεται βίντεο με εξομοίωση της κίνησης του ρομπότ πάνω στο “μονοπάτι” του πανοράματος.

Αλγόριθμος SIFT (Scale Invariant Feature Transform)

Η εύρεση των γεωμετρικών μετασχηματισμών που μετασχηματίζουν την μία εικόνα πάνω στην άλλη προϋποθέτει να γνωρίζουμε αντιστοιχίες σημείων μεταξύ των εικόνων I_N και I_{N+1} . Οι αλγόριθμοι ανίχνευσης αντιστοιχιών που χρησιμοποιούνται πολύ συχνά στην πράξη βασίζονται στις παραγώγους – ακμές της εικόνας (edge ή Corner Detection) . Όμως, τέτοιοι ανιχνευτές (Sobel Prewit, Laplacian operators , Harris Corner Detector) δεν είναι σθεναροί σε κοινές υποβαθμίσεις της εικόνας όπως:

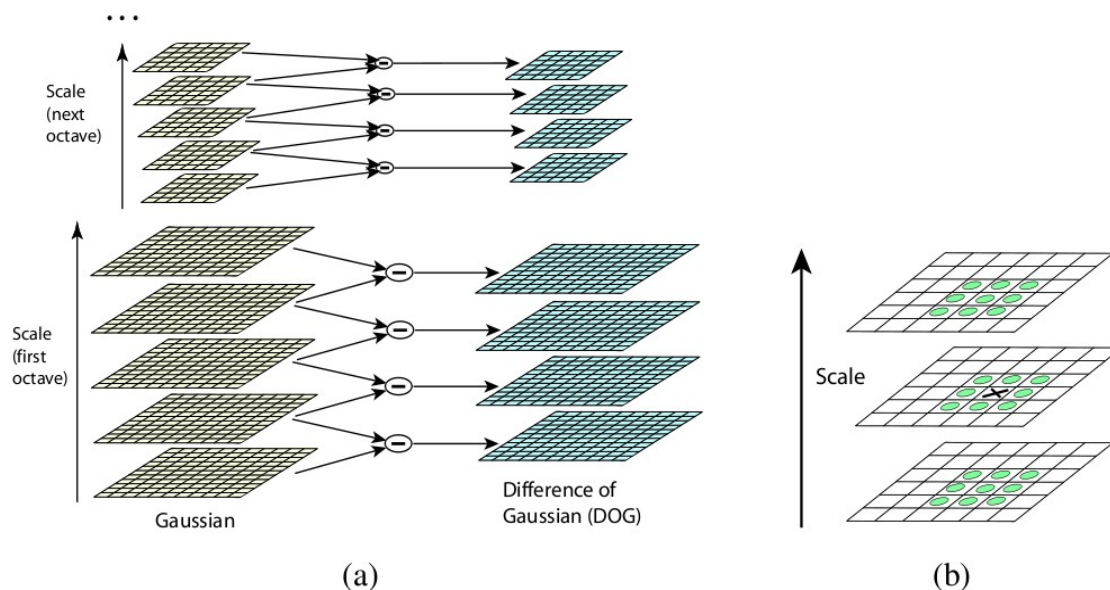
1. παρουσία θορύβου στις εικόνες
2. μη επαρκή φωτισμό ή απότομες μεταβολές του φωτισμού
3. θόλωμα (blurring) από αδυναμία αυτόματης εστίασης της κάμερας

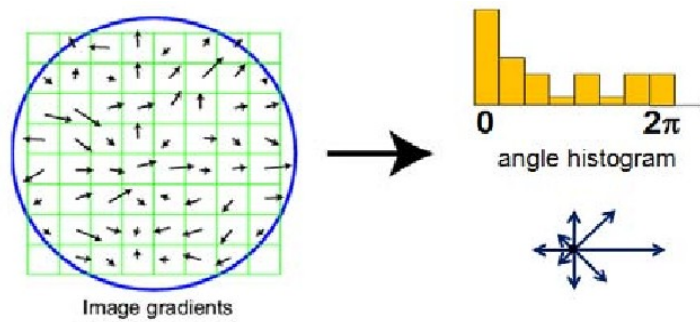
Αντίθετα ο αλγόριθμος SIFT έχει τα παρακάτω ωφέλημα χαρακτηριστικά:

1. Αμεταβλητότητα σε κλιμάκωση, περιστροφή και γενικά affine μετασχηματισμούς και παραμορφώσεις.
2. Αμεταβλητότητα σε αλλαγή φωτισμού και του σημείου λήψης της κάμερας
3. Τα features είναι καλά ορισμένα στο πεδίο του χώρου και της συχνότητας, μειώνοντας την πιθανότητα παραμόρφωσης από θόρυβο ή παρεμβολή.

Ο SIFT βασίζεται στην ιδέα ότι τα features μίας εικόνας μπορούν να περιγραφούν αφαιρώντας την πληροφορία υψηλών συχνοτήτων.

Τα βασικά βήματα του SIFT συνοψίζονται στα παρακάτω σχήματα:





1. Χωρική ανίχνευση “Ακραίων” σημείων - Scale-space extrema detection
Ο αλγόριθμος υπολογίζει την απόκριση κάθε σημείου της εικόνας στο φίλτρο difference-of-Gaussian σε όλες τις δυνατές κλίμακες. Αυτά τα σημεία ενδιαφέροντος παρουσιάζουν αμεταβλητότητα σε κλιμάκωση και προσανατολισμό.
2. Εντοπισμός Σημείων κλειδιών - Keypoint localization
Σε κάθε υποψήφιο σημείο χρησιμοποιείται ένα λεπτομερές μοντέλο ώστε να καθοριστεί η θέση και ο προσανατολισμός του.
3. Εύρεση προσανατολισμού - Orientation assignment
Προσανατολισμός ανατίθεται σε κάθε σημείο βασισμένος στην κατεύθυνση των ακμών στη περιοχή της 26x26 γειτονιάς του.
4. Περιγραφέας Σημείου ενδιαφέροντος - Keypoint descriptor
Υπολογίζονται οι παράγωγοι στη γειτονιά του σημείου και επιλέγεται η ποιο αντιπροσωπευτική κλίμακα αναπαράστασης τους. Τα παραπάνω οδηγούν σε περιγραφή που επιτρέπει σημαντικό επίπεδο τοπικής παραμόρφωσης και αλλαγής στον φωτισμό.

Εύρεση των αντιστοιχιών μεταξύ των Features των εικόνων I_N και I_{N+1}

ο αλγόριθμος SHIFT δέχεται ως είσοδο τις εικόνες I_N και I_{N+1} και επιστρέφει τα feature points μαζί με το διάνυσμα χαρακτηριστικών τους (Keypoint descriptor).

Δύο σημεία μια εικόνας θεωρούμε ότι ταυτίζονται αν:

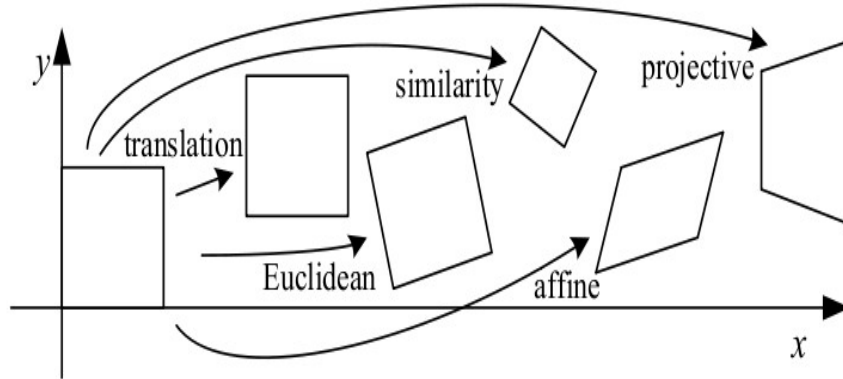
- 1) Η απόσταση των διανυσμάτων χαρακτηριστικών τους είναι μικρότερη k φορές από την απόσταση με κάθε άλλο υποψήφιο feature. Το k είναι αυθαίρετο κατώφλι του σχεδιαστή και επιλέγεται εμπειρικά.
- 2) Ακόμα και μετά το βήμα 1 θα υπάρχουν πολλά mismatches γιατί είναι αναγκαίο να αποκλείσουμε τα matches που είναι outliers.
 - Υπολογίζουμε το μέσο μήκος διανύσματος με αρχή το X_N και τέλος το X_{N+1}
 - Απορρίπτουμε τις αντιστοιχίες σημείων που το διάνυσμα τους έχει πολύ μεγάλο ή πολύ μικρό μήκος σε σχέση με τη μέση τιμή, θέτοντας ένα κατώφλι.

Κριτήριο επιλογής των δύο κατώφλιών:

Στα δύο παραπάνω βήματα επιλέγουμε δύο κατώφλια. Ουσιαστικά προσπαθούμε να απορρίψουμε τους outliers που είναι mismatches. Σε αυτή την προσπάθειά μας όμως κινδυνεύουμε να καταλήξουμε με υπερβολικά λίγα σημεία με συνέπεια να κάνουμε κακή εκτίμηση της τροχιάς. Άρα θα πρέπει ο σχεδιαστής να επιλέξει τα κατώφλια ακολουθώντας αυτά τα δύο κριτήρια.

Υπολογισμός του Γεωμετρικού μετασχηματισμού από αντιστοιχίες σημείων

Στην περίπτωση της 2D εικόνας ο γεωμετρικός μετασχηματισμός κάποιου διανύσματος σε ένα άλλο περιγράφεται από ένα πίνακα 3x3.



Ανάλογα με τον τύπο του μετασχηματισμού που θα εφαρμόσουμε χρειαζόμαστε κατάλληλο αριθμό αντιστοιχιών μη συγγραμμικών σημείων.

Name	Matrix	# D.O.F.	Preserves:	Icon
translation	$\begin{bmatrix} \mathbf{I} & \mathbf{t} \end{bmatrix}_{2 \times 3}$	2	orientation + ...	
rigid (Euclidean)	$\begin{bmatrix} \mathbf{R} & \mathbf{t} \end{bmatrix}_{2 \times 3}$	3	lengths + ...	
similarity	$\begin{bmatrix} s\mathbf{R} & \mathbf{t} \end{bmatrix}_{2 \times 3}$	4	angles + ...	
affine	$\begin{bmatrix} \mathbf{A} \end{bmatrix}_{2 \times 3}$	6	parallelism + ...	
projective	$\begin{bmatrix} \tilde{\mathbf{H}} \end{bmatrix}_{3 \times 3}$	8	straight lines	

Ο ποιο γενικός μετασχηματισμός είναι ο affine ή αυθαίρετος μετασχηματισμός:

$$\mathbf{x}' = \begin{bmatrix} a_{00} & a_{01} & a_{02} \\ a_{10} & a_{11} & a_{12} \end{bmatrix} \tilde{\mathbf{x}}.$$

- Αρχικά υπολογίζουμε τον affine μετασχηματισμό των σημείων της εικόνας I_N στην I_{N+1} και έπειτα χρησιμοποιώντας τον ανάστροφο μετασχηματισμό ταιριάζουμε την I_{N+1} στην I_N .

Αναλυτικότερα τα βήματα που ακολουθήσαμε ήταν:

1. Υπολογίσαμε τον affine μετασχηματισμό των σημείων της εικόνας I_N στην I_{N+1} μέσω της συνάρτησης `Tform=cp2tform(U,X,'affine')` και τους τοποθετήσαμε σε ένα buffer;
2. Για την εικόνα N ο μετασχηματισμός περιστροφής της θα είναι το γινόμενο όλων μετασχηματισμών από τον 1 έως το $N-1$. “Αθροίζουμε τις περιστροφές”.
3. Μετασχηματίζουμε κάθε εικόνα περιστρέφοντας την.
4. Υπολογισμός Μετατόπισης T_x T_y :
 - Λόγω των πολλών πολλαπλασιασμών στους πίνακες μετασχηματισμού δεν μπορούσαμε να υπολογίσουμε αξιόπιστα την μετατόπιση T_x T_y .
 - Για να ξεπεράσουμε αυτό το πρόβλημα εφαρμόζουμε τον αλγόριθμο SHIFT μεταξύ των περιστραμμένων εικόνων.
 - Υπολογίζουμε την μετατόπιση T_x T_y εφαρμόζοντας τον αλγόριθμο RANSAC από το toolbox της PANDEMO.

Ο πίνακας θα έχει μορφή:

1	0	0
0	1	0
T_x	T_y	1

Εφαρμόσαμε τον αλγόριθμο RANSAC από το toolbox της PANDEMO επειδή δεν μπορούσαμε να βρούμε ικανοποιητικό αριθμό αντιστοιχιών μη συγγραμικών σημείων με την απλή μεθοδολογία που προτείναμε στο παραπάνω κεφάλαιο.

5. Η τροχιά είναι τα κεντροειδή των φωτογραφιών όταν αυτές έχουν τοποθετηθεί στο δίοραμα:

$$\begin{aligned} X_{N+1} &= X_N + T_x + X_{\text{centroid}} \\ Y_{N+1} &= Y_N + T_y + Y_{\text{centroid}} \end{aligned}$$

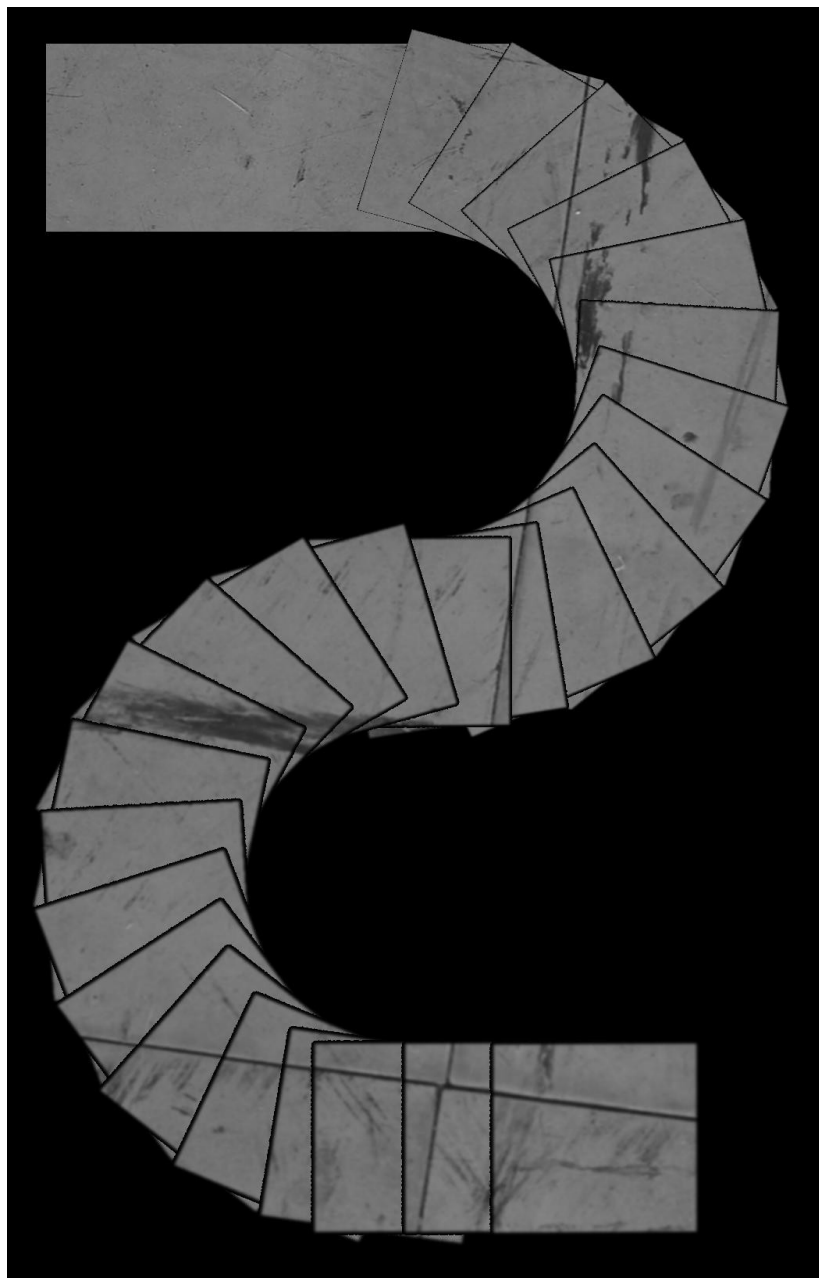
Τα $(X_{\text{centroid}}, Y_{\text{centroid}})$ είναι τα κεντροειδή των περιστραμμένων εικόνων του βήματος 3.

Τα κεντροειδή αποτελούν διορθωτικό όρο καθώς θεωρούμε απριόρι ότι η κάμερα του ρομπότ θα πρέπει να βρήσκεται στο κεντροειδές της εικόνας και κάνουμε την σύμβαση ότι οι αρχικές τιμές της θέσεις θα είναι το κέντρο της πρώτης εικόνας.

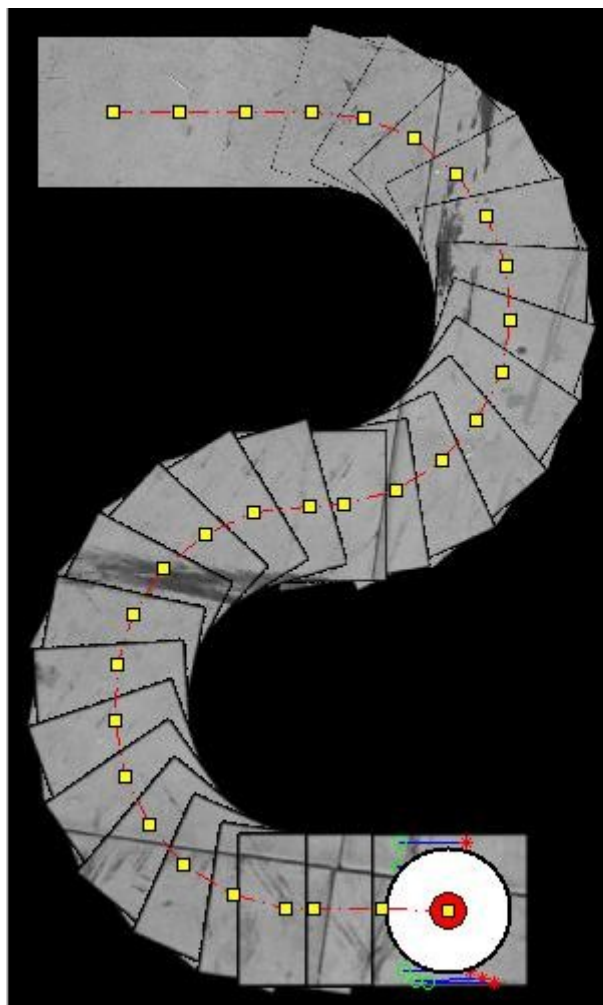
$$(X_0, Y_0) = (128, 128)$$

Αποτελέσματα

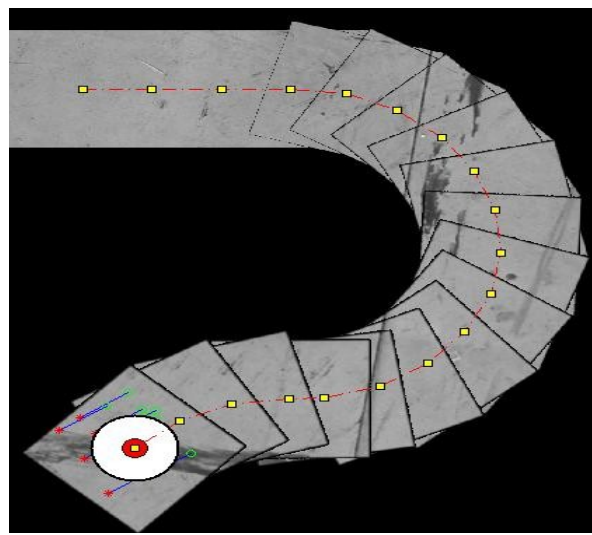
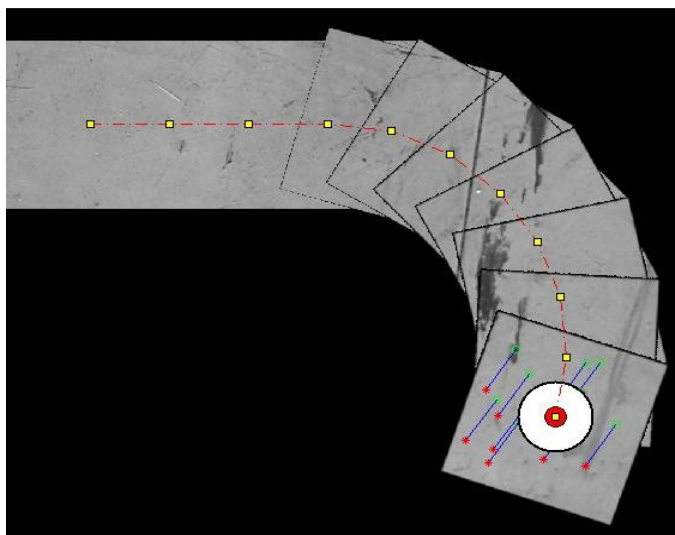
Παρακάτω φαίνεται το δίοραμα που παράγει ο αλγόριθμός μας:



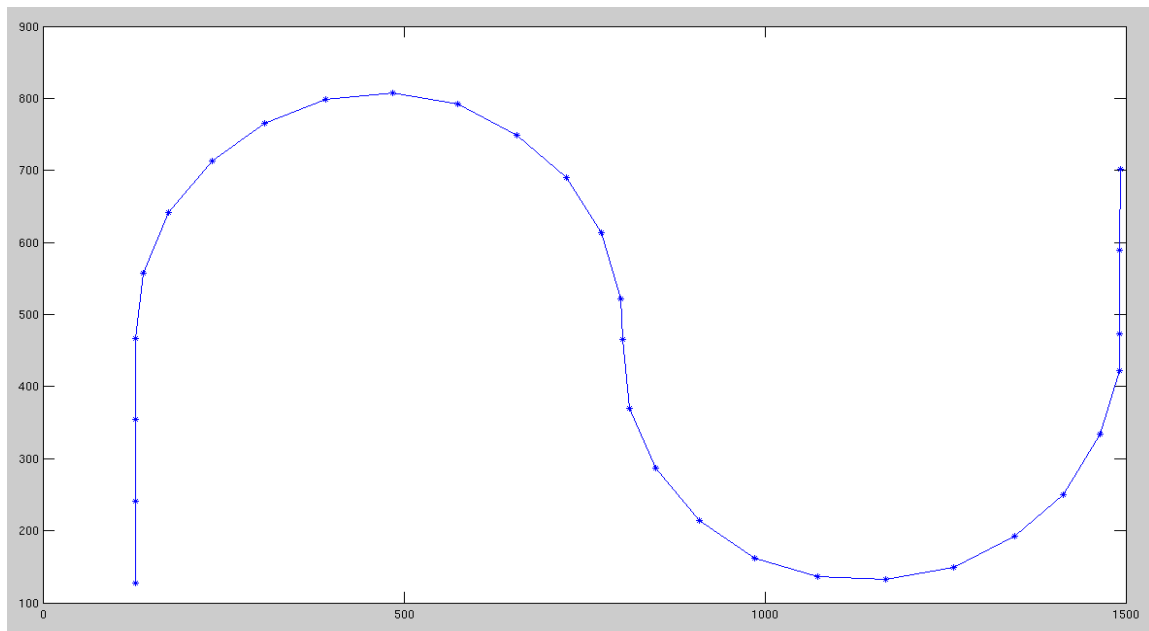
Στην επόμενη εικόνα φαίνεται το διάγραμμα μαζί με το ίχνος της τροχιάς που πραγματοποιεί το Ρομπότ:



Παρακάτω φαίνονται ενδεικτικά frames κατά την εκτέλεση του αλγορίθμου στα οποία απεικονίζεται και η **οπτική ροή με διανύσματα**. Τα διανύσματα προκύπτουν από τα 10 καλύτερα matches ως αποτέλεσμα του RANSAC.



Η τροχιά του Robot φαίνεται στο παρακάτω σχήμα:



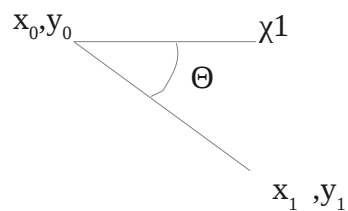
Η τροχιά του Ρομπότ – Η απόσταση μετράται με pixel

Παρατηρούμε ότι το Ρομπότ εκτελεί μια ημιτονοειδή τροχιά.

Στο παραπάνω διάγραμμα η απόσταση μετράται σε pixels. Θα μπορούσαμε να υπολογίσουμε επακριβώς την τροχιά αν γνωρίζαμε τη διάσταση ενός pixel, τον τύπο του αισθητήρα CCD της κάμερας και το εστιακό μήκος με το οποίο δημιουργήθηκε κάθε φωτογραφία.

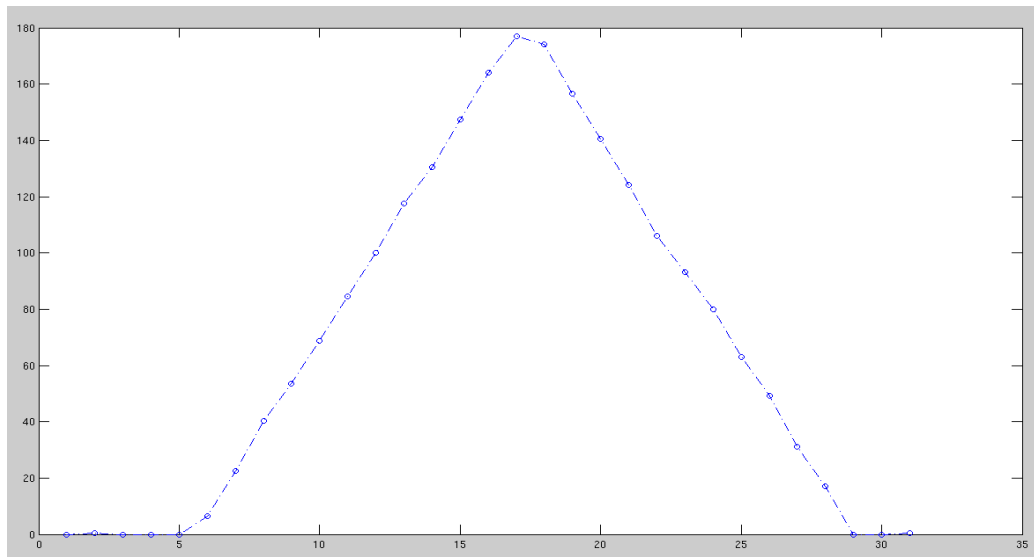
Εκτίμηση Προσανατολισμού του οχήματος

Σύμβαση: θεωρούμε ότι όταν ξεκινάει το ρομπότ ο προσανατολισμός του είναι $\theta = 0$ μοίρες.



Η θ αυξάνεται κατά την ωρολογιακή φορά

Η γωνία προσανατολισμού θ είναι : $\theta = \text{atan2}(x_1 - x_0 / y_1 - y_0) - 45^\circ$



Προσανατολισμός του Ρομπότ - γωνία θ

Βλέποντας το παραπάνω διάγραμμα παρατηρούμε ότι ο προσανατολισμός του Ρομπότ κυμαίνεται μεταξύ 0° και 180° . οι 0° αντιστοιχούν στην αρχική και τελική θέση του Ρομπότ. οι 90° αντιστοιχούν ακριβώς στη μέση της διαδρομής. οι 180° αντιστοιχούν στο δέκατο έβδομο βήμα του αλγορίθμου.

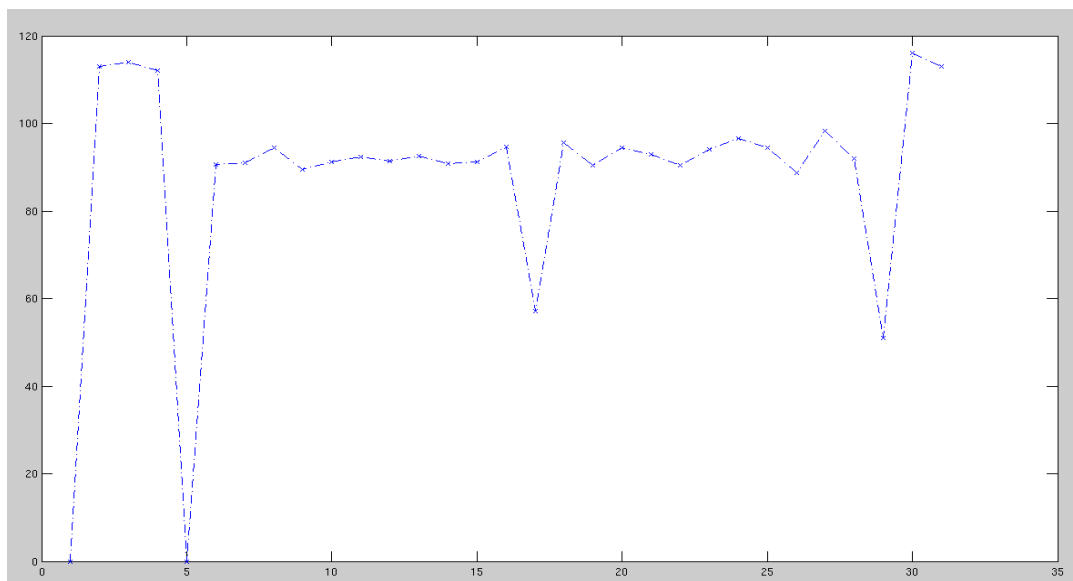
Υπολογισμός Γραμμικής Ταχύτητας

Αφού υπολογίσαμε την τροχιά του ρομπότ μπορούμε να εκτιμήσουμε την Γραμμική Ταχύτητα του οχήματος.

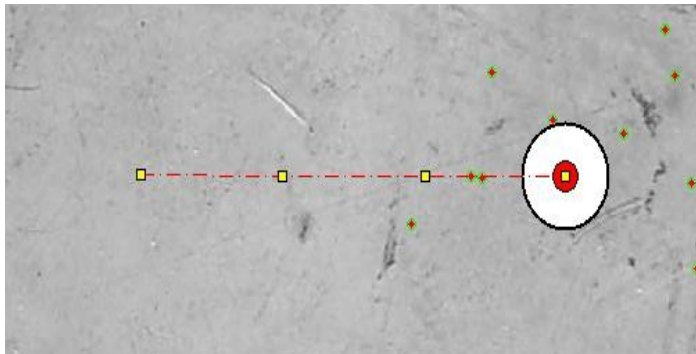
Η ταχύτητα ορίζεται ως: $u = \sqrt{(\frac{dx}{dt})^2 + (\frac{dy}{dt})^2}$

και επειδή η μονάδα χρόνου είναι $dt = 1$ μπορούμε να γράψουμε ότι:

$$u = \sqrt{(X_{N+1} - X_N)^2 + (Y_{N+1} - Y_N)^2}$$



Εκτιμώμενη γραμμική ταχύτητα του robot



Στο Βήμα 5 παρατηρούμε ότι έχουμε **μηδενική Οπτική Ροή**. Το Ρομποτ έχει σταματήσει μηδενική ταχύτητα.

Παρατήρηση: Για να εκτιμήσουμε την ταχύτητα κάναμε την παραδοχή ότι $dt=1$ μονάδα χρόνου.

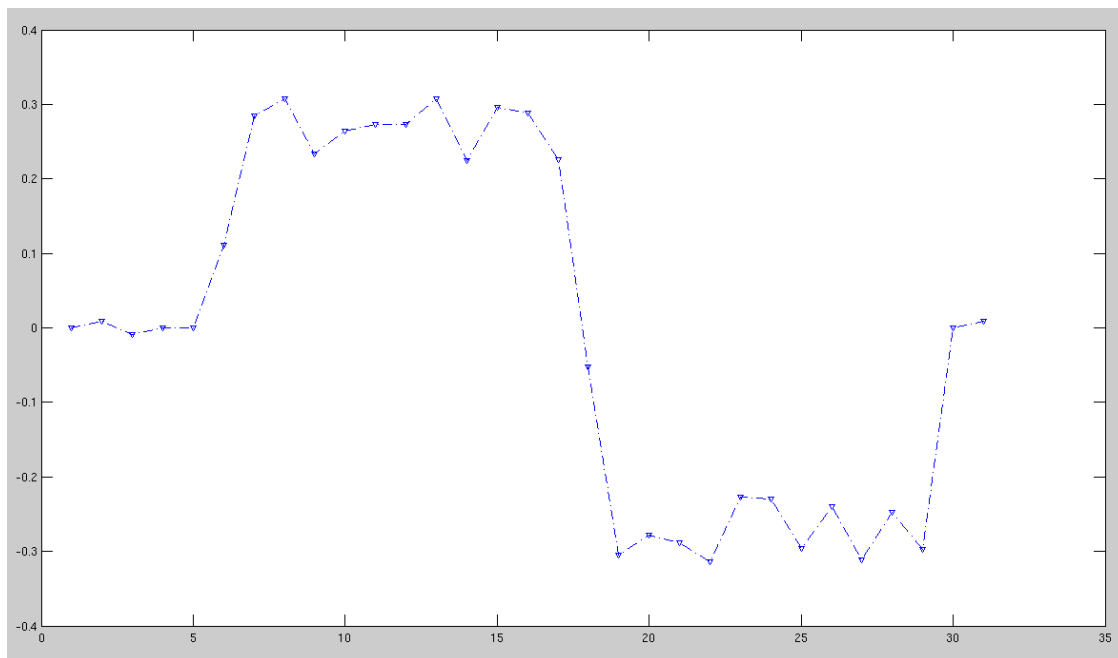
Στην πραγματικότητα, το dt είναι συνάρτηση της συχνότητας δειγματοληψίας f με την οποία η κάμερα λαμβάνει τις εικόνες $dt=1/f$.

Υπολογισμός Γωνιακής Ταχύτητας

$$\omega = d\theta/dt$$

για $dt=1$ μονάδα χρόνου

$$\omega_N = \theta_N - \theta_{N-1}$$



Εικόνα 1: Γωνιακή Ταχύτητα - σε rad/sec

- Μαζί με την αναφορά παραδίδουμε και δύο video που απεικονίζουν:

- 1) Την διαδικασία κατασκευής του πανοράματος
- 2) εξομοίωση της κίνησης του ρομπότ πάνω στο “μονοπάτι” του πανοράματος, με απεικόνιση της οπτικής Ροής.

Βιβλιογραφία:

1. Computer Vision: Algorithms and Applications Richard Szeliski August 18, 2010 draft
2. Image Alignment and Stitching: A Tutorial1 Richard Szeliski Preliminary draft, September 27, 2004 Technical Report MSR-TR-2004-92
3. Object Recognition from Local Scale-Invariant Features David G. Lowe Computer Science Department University of British Columbia Vancouver, B.C., V6T 1Z4, Canada
4. In: Proc. CERN School of Computing, Egmond aan Zee, The Netherlands, 8–21 September, 1996. Scale-space: A framework for handling image structures at multiple scales Tony Lindeberg KTH, S-100 44 Stockholm, Sweden
5. Encyclopedia of Computer Science and Engineering (Benjamin Wah, ed), John Wiley and Sons, Volume IV, pages 2495–2504, Hoboken, New Jersey, 2009.
[dx.doi.org/10.1002/9780470050118.ecse609](https://doi.org/10.1002/9780470050118.ecse609) (Sep 2008) Scale-space Tony Lindeberg School of Computer Science and Communication, KTH (Royal Institute of Technology), SE-100 44 Stockholm, Sweden
6. ΡΟΜΠΟΤΙΚΗ ΟΔΟΜΕΤΡΙΑ ΜΕ ΥΠΟΛΟΓΙΣΤΙΚΗ ΟΡΑΣΗ Ελένη Κελασίδη, Αντώνιος Τζέξ Σχολή Ηλεκτρολόγων μηχανικών και Τεχνολογίας Υπολογιστών Πανεπιστήμιο Πατρών

