

**ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΙΑΣ**  
**ΠΟΛΥΤΕΧΝΙΚΗ ΣΧΟΛΗ**

**ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΗΛΕΚΤΡΟΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ,  
ΤΗΛΕΠΙΚΟΙΝΩΝΙΩΝ ΚΑΙ ΔΙΚΤΥΩΝ**

**«DYNAMIC TEXTURES»**

**ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ**

**ΟΝΟΜΑΤΕΠΩΝΥΜΟ:**

**ΦΙΛΟΘΟΔΩΡΟΣ ΓΕΩΡΓΙΟΣ**

**ΕΠΙΒΛΕΠΟΝΤΕΣ ΚΑΘΗΓΗΤΕΣ:**

**ΜΟΥΣΤΑΚΙΔΗΣ ΓΕΩΡΓΙΟΣ**

**ΜΠΡΙΑΣΣΟΥΛΗ ΑΛΕΞΙΑ**

**ΣΕΠΤΕΜΒΡΙΟΣ 2006**

**ΒΟΛΟΣ**



**ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΙΑΣ  
ΒΙΒΛΙΟΘΗΚΗ & ΚΕΝΤΡΟ ΠΛΗΡΟΦΟΡΗΣΗΣ  
ΕΙΔΙΚΗ ΣΥΛΛΟΓΗ «ΓΚΡΙΖΑ ΒΙΒΛΙΟΓΡΑΦΙΑ»**

Αριθ. Εισ.: 5086/1  
Ημερ. Εισ.: 21-09-2007  
Δωρεά: Συγγραφέα  
Ταξιθετικός Κωδικός: ΠΤ – ΜΗΥΤΔ  
2006  
ΦΙΛ

Στην οικογένειά μου

## **ΕΥΧΑΡΙΣΤΙΕΣ**

Θα ήθελα να ευχαριστήσω θερμά τους επιβλέποντες καθηγητές κ.κ. Μουστακίδη Γεώργιο και Μπριασσούλη Αλεξία για τη συνεργασία, τις χρήσιμες συμβουλές και τη στήριξή τους.

Επίσης θα ήθελα να ευχαριστήσω θερμά την οικογένειά μου για την στήριξη και την υπομονή τους.

# Περιεχόμενα

ΠΡΟΛΟΓΟΣ.....	6
ΜΕΡΟΣ Α.....	7
A.1 Εισαγωγή.....	7
A.2 Ευστάθεια του συστήματος.....	7
A.2α Υπολογισμός της συνάρτησης μεταφοράς.....	8
A.3 Μετασχηματισμός ομοιότητας.....	9
A.4 Ανάλυση Ιδιαζουσών Τιμών (SVD).....	9
A.4α Σχέση svd με eigenvalue decomposition.....	10
A.4β Οικονομική μορφή του svd.....	10
A.5 Ψευδοαντίστροφος.....	11
A.6 Frobenius νόμα.....	11
A.7 Μείωση διάστασης και Principal Components Analysis.....	12
A.7α Περιγραφή του προβλήματος.....	13
A.7β Principal components και svd.....	14
A.8 Παρουσίαση της λύσης.....	15
A.9 Παρουσίαση Αποτελεσμάτων.....	17
A.9α Βίντεο 1.....	17
A.9α.1 Παρουσίαση αποτελεσμάτων αλλάζοντας το μέγεθος του διανύσματος κατάστασης.....	21
A.9α.2 Παρουσίαση χρονικής απόδοσης της προσομοίωσης.....	23
A.9α.2α Χρονική απόδοση του συστήματος κατά το στάδιο της σύνθεσης.....	23
A.9α.2β Χρονική απόδοση του συστήματος κατά το στάδιο της εκμάθησης.....	24
A.9β Βίντεο 2.....	24
A.9β.1 Παρουσίαση αποτελεσμάτων αλλάζοντας το μέγεθος του διανύσματος κατάστασης.....	28
A.10 Συμπύεση.....	30
A.11 Αλλαγή του μεγέθους του συνόλου εκμάθησης και του πλήθους των διανυσμάτων βάσης.....	31
ΜΕΡΟΣ Β.....	33
B.1 Εισαγωγή.....	33
B.2 Υπολογισμός των παραμέτρων.....	33
B.3 Παρουσίαση αποτελεσμάτων.....	34
B.4 Τροποποίηση των πόλων του συστήματος.....	35
B.5 Παρουσίαση των αποτελεσμάτων μετά την τροποποίηση των πόλων.....	37
B.5α Βίντεο 1.....	37
B.5α.1 Χρονική απόδοση του συστήματος κατά το στάδιο της σύνθεσης.....	38
B.5α.2 Χρονική απόδοση του συστήματος κατά το στάδιο της εκμάθησης .....	39
B.5β Βίντεο 2.....	39
ΜΕΡΟΣ Γ.....	41
Γ.1 Εισαγωγή.....	41
Γ.2 Υπολογισμός της μήτρας του συστήματος.....	41

Γ.2α Υπολογισμός του αντιστρόφου.....	43
Γ.3 Μετασχηματισμός Householder.....	44
Γ.3α Γεωμετρική ερμηνεία του μετασχηματισμού Householder .....	45
Γ.3β Εύρεση του διανύσματος Householder .....	46
Γ.4 Υπολογισμός της βάσης του συστήματος.....	47
Γ.4α Παρουσίαση της μεθόδου FDPΜ .....	47
Γ.5 Παρουσίαση αποτελεσμάτων .....	48
Γ.5α Βίντεο 1 .....	51
Γ.5β Βίντεο 2.....	53
Γ.6 Χρονική απόδοση του συστήματος.....	55
<b>Βιβλιογραφία .....</b>	<b>57</b>

## ΠΡΟΛΟΓΟΣ

Στην παρούσα εργασία ασχολούμαστε με μία ειδική κατηγορία βίντεο που ονομάζεται *dynamic texture*. Τα εν λόγω βίντεο παρουσιάζουν κάποιες ιδιότητες τις οποίες δεν έχουν τα συνηθισμένα βίντεο. Συγκεκριμένα *dynamic texture* καλείται μία ακολουθία εικόνων οι οποίες παρουσιάζουν την ιδιότητα της στασιμότητας στο χρόνο. Στην κατηγορία των *dynamic texture* μπορούν να συμπεριληφθούν βίντεο όπως φωτιά, κύματα θάλασσας, καπνός, σιντριβάνι κ.λ.π.

Υπάρχουν διάφορες εφαρμογές, όπως βιντεοπαιχνίδια και *screensavers*, στις οποίες θα θέλαμε τα παραπάνω βίντεο να διαρκούν για όσο διάστημα επιθυμεί ο χρήστης. Ωστόσο δεν θα θέλαμε το τελικό βίντεο να αποτελείται από μία ακολουθία καρέ η οποία θα επαναλαμβάνεται πολλές φορές, αλλά από καρέ τα οποία είναι διαφορετικά μεταξύ τους και τα οποία δεν παρουσιάζουν ασυνέχειες. Για να επιτευχθεί κάτι τέτοιο χρησιμοποιούμε ένα μαθηματικό μοντέλο το οποίο αρχικά χρησιμοποιεί μία ακολουθία από καρέ για να μάθει την κίνηση και κατόπιν συνθέτει νέα καρέ τα οποία μοιάζουν με τα αρχικά αλλά είναι διαφορετικά.

## ΜΕΡΟΣ Α

### A.1 Εισαγωγή

Οι Gianfranco Doretto, Alessandro Chiuso, Ying N.Wu και Stefano Soatto υιοθέτησαν ένα γραμμικό δυναμικό σύστημα για να αναπαραστήσουν και να συνθέσουν dynamic textures. Το μοντέλο που προτείνουν αναπαρίσταται στο χώρο κατάστασης και έχει ως ακολούθως:

$$\begin{aligned}x(n+1) &= Ax(n) + Bu(n+1) \\ y(n) &= Cx(n) + w(n)\end{aligned}\quad (A.1)$$

Η αναλυτική περιγραφή του προτεινόμενου μοντέλου είναι η εξής:

Έστω  $k \times l$  η ανάλυση ενός καρέ και  $m$  ο αριθμός εικονοστοιχείων της εικόνας, δηλαδή  $m = k \times l$ . Στο μοντέλο που περιγράφουμε κάθε καρέ αναπαρίσταται ως  $I(t)$  και δημιουργείται βάζοντας τις  $l$  στήλες της εικόνας τη μία κάτω από την άλλη. Συνεπώς το  $I(t)$  είναι ένα διάνυσμα στήλη με τόσα στοιχεία όσα τα εικονοστοιχεία ενός καρέ. Έτσι  $I(t) \in \mathfrak{R}^m$ . Επίσης έχουμε μία ακολουθία  $\{I(t)\}$  η οποία αποτελείται από  $\tau$  καρέ και συνιστά την ακολουθία από την οποία το μαθηματικό μοντέλο θα υπολογίσει τις παραμέτρους κατά τη διαδικασία μάθησης. Το  $A \in \mathfrak{R}^{n \times n}$  είναι ο πίνακας του συστήματος, το διάνυσμα  $x(t) \in \mathfrak{R}^n$  συνιστά τις μεταβλητές κατάστασης, το  $u(t) \in \mathfrak{R}^m$  είναι μία ανεξάρτητη και όμοια κατανομημένη ακολουθία με κατανομή  $q()$ ,  $B \in \mathfrak{R}^{n \times m}$ ,  $C \in \mathfrak{R}^{m \times n}$  είναι ο πίνακας εξόδου και  $w(t) \in \mathfrak{R}^m$  είναι μία ανεξάρτητη και όμοια κατανομημένη ακολουθία με γνωστή κατανομή  $p()$ .

Στη γενική περίπτωση

$$\begin{aligned}x(n) &= \sum_{i=1}^k A_i x(n-i) + Bu(n) \\ y(n) &= f(x(n)) + w(n)\end{aligned}\quad (A.2)$$

Στο αρχικό μοντέλο  $f(x(n)) = \sum_{i=1}^n \Theta_i x_i(n) = Cx(n)$ , όπου  $\Theta_i \in \mathfrak{R}^m$  μπορεί να είναι μία ορθοκανονική βάση, ένα σύνολο από principal components κλπ. Η επιλογή της  $f(x(t))$  μπορεί να θεωρηθεί σαν ένα βήμα για τη μείωση της διάστασης.

### A.2 Ευστάθεια του συστήματος

Όπως ήδη έχω αναφέρει στόχος μας είναι να συνθέτουμε καρέ για άπειρο χρονικό διάστημα. Για να καταστεί εφικτό κάτι τέτοιο θα πρέπει το μοντέλο μας να είναι ευσταθές, θα πρέπει δηλαδή οι πόλοι του συστήματος να βρίσκονται μέσα στο



μοναδιαίο κύκλο. Έτσι θα πρέπει να υπολογίσουμε με τη βοήθεια του μετασχηματισμού  $Z$  τη συνάρτηση μεταφοράς του συστήματος

### A.2α Υπολογισμός της συνάρτησης μεταφοράς

Αρχικά παίρνουμε το μετασχηματισμό  $Z$  της καταστατικής εξίσωσης και λύνουμε ως προς  $X(z)$ :

$$\begin{aligned} Z\{x(n)\} &= Z\{Ax(n-1) + Bu(n)\} \Rightarrow \\ X(z) &= Az^{-1}X(z) + BU(z) \Rightarrow \\ X(z) - Az^{-1}X(z) &= BU(z) \Rightarrow \\ [I - Az^{-1}]X(z) &= BU(z) \Rightarrow \\ X(z) &= [I - Az^{-1}]^{-1} BU(z) \quad (A.3) \end{aligned}$$

Έπειτα παίρνουμε το μετασχηματισμό  $Z$  της εξίσωσης εξόδου και αντικαθιστούμε το  $X(z)$  με το παραπάνω αποτέλεσμα:

$$\begin{aligned} Z\{y(n)\} &= Z\{Cx(n)\} \Rightarrow \\ Y(z) &= CX(z) \Rightarrow \\ Y(z) &= C[I - Az^{-1}]^{-1} BU(z) \Rightarrow \\ H(z) &= \frac{Y(z)}{U(z)} = C[I - Az^{-1}]^{-1} B \quad (A.4) \end{aligned}$$

Χρησιμοποιώντας τον ορισμό του αντιστρόφου ενός πίνακα το τελευταίο αποτέλεσμα μπορεί να γραφτεί ως εξής:

$$\begin{aligned} H(z) &= C[I - Az^{-1}]^{-1} B \Rightarrow \\ H(z) &= Cz[zI - A]^{-1} B \Rightarrow \\ H(z) &= Cz \frac{\text{adj}(zI - A)}{|zI - A|} B \quad (A.5) \end{aligned}$$

όπου  $\text{adj}(zI - A)$  είναι το αλγεβρικό συμπλήρωμα του τετραγωνικού πίνακα  $zI - A$  και  $|zI - A|$  η ορίζουσα του πίνακα  $zI - A$ . Γνωρίζουμε ότι οι πόλοι ενός γραμμικού συστήματος είναι εκείνες οι τιμές του  $z$  στις οποίες η συνάρτηση μεταφοράς γίνεται άπειρη. Από την τελευταία σχέση παρατηρούμε ότι η συνάρτηση μεταφοράς απειρίζεται για εκείνα τα  $z$  που μηδενίζουν την ορίζουσα του πίνακα  $zI - A$ , οι πόλοι δηλαδή του συστήματος είναι οι ρίζες του χαρακτηριστικού πολυωνύμου  $D(z) = |zI - A|$  ή ισοδύναμα οι ιδιοτιμές του πίνακα  $A$ .

### A.3 Μετασχηματισμός ομοιότητας

Ως μετασχηματισμός ομοιότητας ενός συστήματος που περιγράφεται στο χώρο κατάστασης ορίζεται:

$$x(n) = Tx'(n) \quad (A.6)$$

όπου  $T$  οποιοσδήποτε αντιστρέψιμος πίνακας και  $x'(n)$  το νέο διάνυσμα κατάστασης.

Αντικαθιστώντας την παραπάνω σχέση στο μοντέλο που περιγράφει ένα σύστημα στο χώρο κατάστασης παίρνουμε:

$$\begin{aligned} Tx'(n+1) &= A \cdot T \cdot x'(n) + B \cdot U(n) \\ y(n) &= C \cdot T \cdot x'(n) \end{aligned} \quad (A.7)$$

Πολλαπλασιάζοντας την εξίσωση κατάστασης με τον αντίστροφο του πίνακα  $T$  παίρνουμε:

$$\begin{aligned} x'(n+1) &= T^{-1} A \cdot T \cdot x'(n) + T^{-1} B \cdot U(n) \\ y(n) &= C \cdot T \cdot x'(n) \end{aligned} \quad (A.8)$$

Έτσι οι πίνακες του συστήματός μας μετασχηματίζονται ως εξής:

$$\begin{aligned} A' &= T^{-1} \cdot A \cdot T \\ B' &= T^{-1} \cdot B \\ C' &= C \cdot T \end{aligned} \quad (A.9)$$

Αν εφαρμόσουμε ένα μετασχηματισμό ομοιότητας η συνάρτηση μεταφοράς του συστήματος παραμένει η ίδια:

$$\begin{aligned} H(z) &= C' \cdot z \cdot (zI - A')^{-1} \cdot B' = \\ &= C \cdot T \cdot z \cdot (zI - T^{-1} \cdot A \cdot T)^{-1} \cdot T^{-1} \cdot B = \\ &= C \cdot z \cdot [T(zI - T^{-1} \cdot A \cdot T) \cdot T^{-1}]^{-1} \cdot B = \\ &= C \cdot z \cdot [(zT - A \cdot T) \cdot T^{-1}]^{-1} \cdot B = \\ &= C \cdot z \cdot (zI - A)^{-1} \cdot B \end{aligned} \quad (A.10)$$

### A.4 Ανάλυση Ιδιαζουσών Τιμών (SVD)

Για κάθε πραγματικό πίνακα  $A$  διάστασης  $n \times m$  υπάρχει μία παραγοντοποίηση της μορφής  $A = U \Sigma V^T = \sum_{k=1}^K \sigma_k u_k v_k$ , όπου  $K = \min(m, n)$ ,  $u_k$  η  $k$  στήλη του πίνακα  $U$ ,  $v_k$  η  $k$  γραμμή του πίνακα  $V$  και  $\sigma_k$  το  $(k, k)$  στοιχείο του διαγώνιου πίνακα  $\Sigma$ . Τα στοιχεία των πινάκων μπορούν να αναδιαταχθούν έτσι ώστε

οι τιμές τις διαγωνίου του πίνακα  $\Sigma$ , που καλούνται ιδιάζουσες τιμές, να είναι διατεταγμένες κατά φθίνουσα σειρά, δηλαδή  $\sigma_1 \geq \dots \geq \sigma_k \geq 0$ .

Οι πίνακες  $U$  και  $V$  έχουν διάσταση  $n \times n$  και  $m \times m$  αντίστοιχα και είναι ορθογώνιοι και κανονικοί, δηλαδή ισχύουν οι παρακάτω σχέσεις:

$$u_a \cdot u_b = \begin{cases} 1 & , \quad a = b \\ 0 & , \quad a \neq b \end{cases} \quad a, b = 1, \dots, n \quad (A.11)$$

και

$$v_a \cdot v_b = \begin{cases} 1 & , \quad a = b \\ 0 & , \quad a \neq b \end{cases} \quad a, b = 1, \dots, m \quad (A.12)$$

και υπό μορφή πινάκων:

$$U^T \cdot U = I_n \quad \text{και} \quad V^T \cdot V = I_m$$

Οι στήλες του πίνακα  $U$  ονομάζονται αριστερά ιδιάζοντα διανύσματα και οι στήλες του πίνακα  $V$  δεξιά ιδιάζοντα διανύσματα.

#### A.4α Σχέση svd με eigenvalue decomposition

Αν και το singular value decomposition μπορεί να εφαρμοστεί σε οποιοδήποτε πίνακα σε αντίθεση με το eigenvalue decomposition που εφαρμόζεται σε τετραγωνικούς πίνακες οι δύο αποδομήσεις συσχετίζονται.

Έστω ένας πίνακας  $A = USV^T$ . Πολλαπλασιάζοντας με τον ανάστροφό του παίρνουμε:

$$AA^T = USV^T V S U^T = U \Sigma^2 U^T \Rightarrow AA^T U = U \Sigma^2 U^T U \Rightarrow AA^T U = U \Sigma^2 \quad (A.13)$$

$$A^T A = V S U^T U S V^T = V \Sigma^2 V^T \Rightarrow A^T A V = V \Sigma^2 V^T V \Rightarrow A^T A V = V \Sigma^2 \quad (A.14)$$

Από τις παραπάνω σχέσεις οι στήλες του  $U$  και του  $V$  αποτελούν τα ιδιοδιανύσματα του  $AA^T$  και του  $A^T A$  αντίστοιχα με ιδιοτιμές το τετράγωνο των ιδιαζουσών τιμών.

#### A.4β Οικονομική μορφή του svd

Έστω ένας πίνακας  $A = USV^T$  διάστασης  $n \times m$ . Αν  $n > m$  η αποδόμηση ιδιαζουσών τιμών γράφεται:

$$\begin{bmatrix} u_{11} & \dots & u_{1n} \\ \vdots & \ddots & \vdots \\ u_{n1} & \dots & u_{nn} \end{bmatrix} \begin{bmatrix} \sigma_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \sigma_m \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 \end{bmatrix} \begin{bmatrix} v_{11} & \dots & v_{1m} \\ \vdots & \ddots & \vdots \\ u_{m1} & \dots & u_{mm} \end{bmatrix} = \begin{bmatrix} u_{11} & \dots & u_{1m} \\ \vdots & \ddots & \vdots \\ u_{n1} & \dots & u_{nm} \end{bmatrix} \begin{bmatrix} \sigma_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \sigma_m \end{bmatrix} \begin{bmatrix} v_{11} & \dots & v_{1m} \\ \vdots & \ddots & \vdots \\ u_{m1} & \dots & u_{mm} \end{bmatrix}$$

δηλαδή κρατάμε τις  $m$  πρώτες στήλες και γραμμές των πινάκων  $U$  και  $\Sigma$  αντίστοιχα.

Αν  $n < m$ , τότε έχουμε:

$$\begin{bmatrix} u_{11} & \dots & u_{1n} \\ \vdots & \ddots & \vdots \\ u_{n1} & \dots & u_{nn} \end{bmatrix} \begin{bmatrix} \sigma_1 & \dots & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \dots & \sigma_n & \dots & 0 \end{bmatrix} \begin{bmatrix} v_{11} & \dots & v_{1m} \\ \vdots & \ddots & \vdots \\ u_{m1} & \dots & u_{mm} \end{bmatrix} = \begin{bmatrix} u_{11} & \dots & u_{1n} \\ \vdots & \ddots & \vdots \\ u_{n1} & \dots & u_{nn} \end{bmatrix} \begin{bmatrix} \sigma_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \sigma_n \end{bmatrix} \begin{bmatrix} v_{11} & \dots \\ \vdots & \ddots \\ u_{n1} & \dots \end{bmatrix}$$

δηλαδή κρατάμε τις  $n$  πρώτες γραμμές και στήλες των πινάκων  $V$  και  $\Sigma$  αντίστοιχα.

## A.5 Ψευδοαντίστροφος

Γνωρίζουμε ότι για κάθε μη ιδιόμορφο τετραγωνικό πίνακα  $A$  υπάρχει ένας πίνακας  $B$  ίδιας διάστασης τέτοιος ώστε το γινόμενο τους να είναι ο μοναδιαίος πίνακας, δηλαδή  $A \cdot B = B \cdot A = I$ . Ο πίνακας  $B$  ονομάζεται αντίστροφος του  $A$ .

Ο ψευδοαντίστροφος επεκτείνει την έννοια του αντιστρόφου και για μη τετραγωνικούς πίνακες. Έστω πίνακας  $A$  διάστασης  $m \times n$ , με  $m > n$  και  $\text{rank}(A) = n$ . Ο  $n \times n$  πίνακας  $A^T A$  είναι τάξης  $n$  και συνεπώς υπάρχει ο αντίστροφός του  $(A^T A)^{-1}$ . Πολλαπλασιάζοντας το τελευταίο αποτέλεσμα με  $A^T$  παίρνουμε τον ψευδοαντίστροφο  $(A^T A)^{-1} A^T$  του πίνακα  $A$ . Ο ψευδοαντίστροφος που μόλις περιγράψαμε είναι αριστερός αντίστροφος του  $A$  καθώς  $(A^T A)^{-1} A^T A = (A^T A)^{-1} (A^T A) = I_n$ , αλλά όχι δεξιός γιατί η τάξη του  $A(A^T A)^{-1} A^T$  είναι  $n$  κάτι που σημαίνει ότι δεν μπορεί να ισούται με τον μοναδιαίο πίνακα  $I_m$ , του οποίου η τάξη είναι  $m$ . Αντίθετα αν  $m < n$ , τότε ο πίνακας  $A$  θα έχει μόνο δεξιό αντίστροφο που θα είναι ο  $A^T (A A^T)^{-1}$ . Όταν ο πίνακας  $A$  είναι τετραγωνικός ο ψευδοαντίστροφός του είναι ίδιος με τον αντίστρόφό του.

## A.6 Frobenius νόρμα

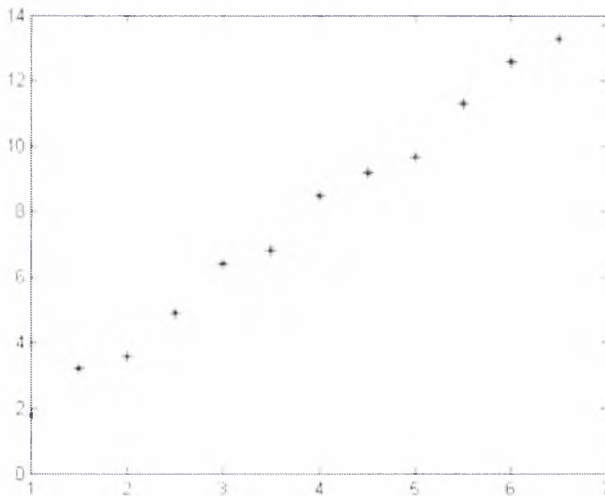
Η frobenius νόρμα ενός πίνακα  $A$  διάστασης  $m \times n$  είναι η τετραγωνική ρίζα του αθροίσματος των τετραγώνων των απολύτων τιμών των στοιχείων του και

συμβολίζεται ως  $\|A\|_F$ . Δηλαδή  $\|A\|_F = \sqrt{\sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2}$ . Επίσης η frobenius νόρμα

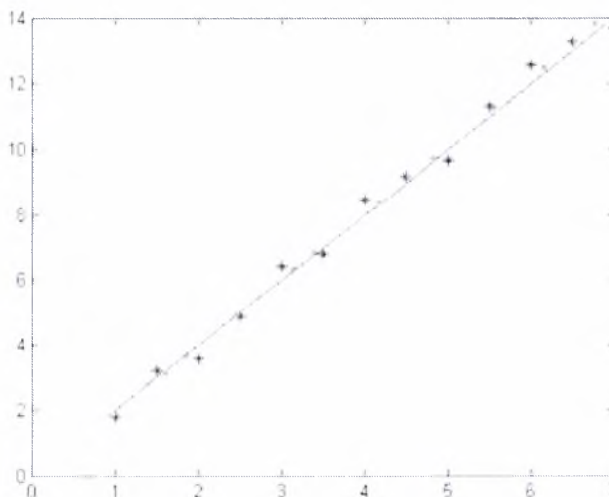
ισούται με  $\|A\|_F = \sqrt{\text{trace}(A \cdot A^T)} = \sqrt{\text{trace}(A^T \cdot A)}$ , όπου ως  $\text{trace}(A)$  ορίζεται το άθροισμα των στοιχείων της κύριας διαγωνίου του πίνακα  $A$ .

## A.7 Μείωση διάστασης και Principal Components Analysis

Όπως ήδη έχουμε αναφέρει την εξίσωση εξόδου του συστήματός μας μπορούμε να τη δούμε ως ένα βήμα για την μείωση της διάστασης. Η μείωση διάστασης είναι επιθυμητή σε κάθε εφαρμογή καθώς για παράδειγμα με τον τρόπο αυτό μειώνεται η χρονική πολυπλοκότητα του προβλήματος. Ακολουθεί ένα διάγραμμα στο οποίο τα δεδομένα απεικονίζονται στον  $\mathbb{R}^2$  (χρησιμοποιώντας δηλαδή δύο διανύσματα βάσης).



Ωστόσο τα δεδομένα αυτά μπορούν να περιγραφούν χρησιμοποιώντας ένα διάνυσμα βάσης έχοντας μια μικρή απώλεια πληροφορίας. Ακολουθεί το διάγραμμα με τα δεδομένα που παίρνουμε χρησιμοποιώντας ένα διάνυσμα βάσης.



Υπάρχουν διάφορες τεχνικές οι οποίες μπορούν να επιτύχουν μείωση της διάστασης ενός προβλήματος, μία εκ των οποίων είναι η Principal Components Analysis (PCA).

### A.7α Περιγραφή του προβλήματος

Έστω  $y = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} \in \mathfrak{R}^m$  το διάνυσμα των δεδομένων μας. Από την Γραμμική

Άλγεβρα γνωρίζουμε ότι το  $y$  μπορεί να γραφεί ως  $y = z_1 \cdot u_1 + \dots + z_m \cdot u_m$ , όπου  $z_i$

βαθμωτό μέγεθος και  $u_i = \begin{bmatrix} u_{1i} \\ \vdots \\ u_{mi} \end{bmatrix} \in \mathfrak{R}^m$ ,  $i = 1, \dots, d$ , ένα σύνολο από  $m$  ορθογώνια

διανύσματα που αποτελούν μία βάση. Το ζητούμενο είναι βρούμε ένα σύνολο από  $d < m$  ορθοκανονικά διανύσματα βάσης  $\hat{u}_i \in \mathfrak{R}^m$ , τέτοια ώστε να ελαχιστοποιείται η  $\|y - \hat{y}\|_F^2$ , όπου  $\hat{y} = \hat{z}_1 \cdot \hat{u}_1 + \dots + \hat{z}_d \cdot \hat{u}_d$  ή να μεγιστοποιείται το  $Var(\hat{u} \cdot y)$ , όπου  $\hat{u} = [\hat{u}_1 \dots \hat{u}_d]$ . Στην τελευταία περίπτωση θέλουμε να βρούμε ένα πίνακα από  $d$  διανύσματα που θα μεγιστοποιεί τη διασπορά έτσι ώστε να είναι πιο εμφανής η διαφορά μεταξύ των δειγμάτων. Από την Γραμμική Άλγεβρα γνωρίζουμε ότι

$$\begin{aligned} \hat{u}_i \cdot y &= \|\hat{u}_i\| \|y\| \cos(\text{angle}(\hat{u}_i, y)) = \\ &= \|\hat{u}_i\| \|y\| \frac{\text{proj}(y \rightarrow \hat{u}_i)}{\|y\|} \end{aligned} \quad (A.15)$$

καθώς  $\text{angle}(\hat{u}_i, y) = \frac{\text{proj}(y \rightarrow \hat{u}_i)}{\|y\|}$ , όπου  $\text{angle}(\hat{u}_i, y)$  είναι η γωνία μεταξύ των διανυσμάτων  $\hat{u}_i$  και  $y$  και  $\text{proj}(y \rightarrow \hat{u}_i)$  η προβολή του  $y$  στο  $\hat{u}_i$ . Αφού τα  $\hat{u}_i$  είναι κανονικά έχουμε ότι  $\|\hat{u}_i\| = 1$ . Έτσι η παραπάνω σχέση γίνεται

$$\hat{u}_i \cdot y = \text{proj}(y \rightarrow \hat{u}_i) \quad (A.16)$$

Συνεπώς, και ισοδύναμα με αυτό που αναφέραμε παραπάνω, θέλουμε να βρούμε διανύσματα  $\hat{u}_i$  στα οποία όταν προβάλουμε το  $y$  να μεγιστοποιείται η διασπορά του  $z = \hat{u}_i \cdot y$ . Παρακάτω παρουσιάζουμε την λύση του προβλήματος μεγιστοποίησης. Αρχικά έχουμε:

$$\begin{aligned} Var(u^T \cdot y) &= E_y \left[ (u^T \cdot y - E_y[u^T \cdot y])^2 \right] = \\ &= E_y \left[ (u^T \cdot y - u^T \cdot E_y[y])^2 \right] = \\ &= E_y \left[ (u^T \cdot y - u^T \cdot \mu_y)^2 \right] = \\ &= E_y \left[ (u^T \cdot (y - \mu_y))^2 \right] = \\ &= u^T \cdot E_y[(y - \mu_y)^2] \cdot u \Rightarrow \\ &\Rightarrow Var(u^T \cdot y) = u^T \cdot \Sigma \cdot u \end{aligned} \quad (A.17)$$

όπου  $\Sigma = E_y \left[ (y - \mu_y)^T (y - \mu_y) \right]$  είναι η μήτρα συνδιασποράς.

Έτσι έχουμε το εξής πρόβλημα:

$$\begin{aligned} \max \quad & u^T \cdot \Sigma \cdot u \\ \text{s.t.:} \quad & u^T \cdot u = 1 \end{aligned} \quad (A.18)$$

Για να το επιλύσουμε θα χρησιμοποιήσουμε τη μέθοδο των πολλαπλασιαστών Lagrange:

$$\begin{aligned} L(u, \lambda) = & u^T \cdot \Sigma \cdot u - \lambda(u^T u - 1) \\ & u^T \cdot u = 1 \end{aligned} \quad (A.19)$$

Παραγωγίζοντας το  $L(u, \lambda)$  ως προς  $u$  και θέτοντάς το ίσο με το μηδέν παίρνουμε:

$$2\Sigma \cdot u - 2\lambda u = 0 \Rightarrow 2\Sigma \cdot u = 2\lambda u \Rightarrow \Sigma \cdot u = \lambda u \quad (A.20)$$

Συνεπώς τα βέλτιστα διανύσματα  $u_i$  είναι τα ιδιοδιανύσματα της μήτρας συνδιασποράς. Τώρα θέλουμε να βρούμε  $d$  από τα ιδιοδιανύσματα που μεγιστοποιούν το άθροισμα:

$$\max \sum_{i=1}^d \hat{u}_i^T \cdot \Sigma \cdot \hat{u}_i$$

Αντικαθιστώντας  $\Sigma \cdot \hat{u}_i = \lambda_i \hat{u}_i$  παίρνουμε:

$$\max \sum_{i=1}^d \hat{u}_i^T \cdot \lambda_i \cdot \hat{u}_i \rightarrow \max \sum_{i=1}^d \lambda_i \hat{u}_i^T \cdot \hat{u}_i \rightarrow \max \sum_{i=1}^d \lambda_i \quad (A.21)$$

Έτσι τα «καλύτερα»  $d$  ιδιοδιανύσματα είναι αυτά που αντιστοιχούν στις  $d$  μεγαλύτερες ιδιοτιμές. Τα διανύσματα αυτά ονομάζονται principal components.

## A.7β Principal components και svd

Όπως αναφέραμε προηγουμένως όταν θέλουμε να περιγράψουμε ένα διάνυσμα με  $d$  διανύσματα βάσης δημιουργούμε τη μήτρα συνδιασποράς των δεδομένων μας και επιλέγουμε τα  $d$  ιδιοδιανύσματα που αντιστοιχούν στις μεγαλύτερες ιδιοτιμές. Επίσης από την ανάλυση ιδιαιζουσών τιμών γνωρίζουμε ότι τα ιδιοδιανύσματα ενός πίνακα  $A \cdot A^T$  είναι ίδια με τα αριστερά ιδιάζοντα διανύσματα του πίνακα  $A$ , δηλαδή με τις στήλες του πίνακα  $U$ , με  $A = U \cdot \Sigma \cdot V^T$ . Έτσι αν κάνουμε svd στον πίνακα  $(y - \bar{y})$ , τότε τα  $d$  διανύσματα βάσης θα είναι οι πρώτες  $d$  στήλες του αριστερού ιδιάζοντος πίνακα.

## A.8 Παρουσίαση της λύσης

Το σύστημα που παρουσιάσαμε στο πρώτο εδάφιο μπορεί να ξαναγραφεί ως :

$$\begin{aligned} x(n+1) &= Ax(n) + v(n+1) \\ y(n) &= Cx(n) + w(n) \end{aligned} \quad (A.22)$$

όπου οι στοχαστικές διαδικασίες  $v(n+1)$  και  $w(n)$  αποτελούν λευκό γκαουσιανό θόρυβο μηδενικής μέσης τιμής με μήτρες συνδιασποράς  $Q$  και  $R$  αντίστοιχα, δηλαδή:

$$\begin{aligned} v(n) &\sim N(0, Q) \\ w(n) &\sim N(0, R) \end{aligned}$$

Στο μοντέλο (A.1) ισχύει ότι  $B \cdot B^T = Q$  και  $u(n) \sim N(0, I_{mv})$  και έτσι τα δύο μοντέλα είναι ίδια.

Έχουμε πει ότι το κάθε καρέ αναπαρίσταται ως ένα  $m$ -διάστατο διάνυσμα και ότι αρχικά έχουμε  $\tau$  καρέ. Έτσι δημιουργούμε τον πίνακα  $Y_1^\tau = [y(1) \dots y(\tau)] \in \mathfrak{R}^{m \times \tau}$ . Δοθέντος λοιπόν του πίνακα καλούμαστε να υπολογίσουμε όλες τις παραμέτρους, δηλαδή τα  $A, B, C, Q$ . Επίσης δημιουργούμε τον πίνακα  $X_1^\tau = [x(1) \dots x(\tau)] \in \mathfrak{R}^{n \times \tau}$  που θα περιέχει τις μεταβλητές κατάστασης για τις  $\tau$  χρονικές τιμές και τον πίνακα  $W_1^\tau = [w(1) \dots w(\tau)] \in \mathfrak{R}^{m \times \tau}$ . Η εξίσωση εξόδου γίνεται:

$$Y_1^\tau = C \cdot X_1^\tau + W_1^\tau \quad (A.23)$$

και θέλουμε να υπολογίσουμε τα  $C, X$  έτσι ώστε:

$$[C, X] = \arg \min \|W_1^\tau\|_F \quad (A.24)$$

Από την ανάλυση που κάναμε στο εδάφιο περί μείωσης της διάστασης είδαμε ότι τα «καλύτερα»  $n$  διανύσματα βάσης είναι τα πρώτα  $n$  αριστερά ιδιάζοντα διανύσματα του πίνακα  $Y_1^\tau$ , με  $Y_1^\tau = U \cdot \Sigma \cdot V^T$ . Έτσι παίρνουμε  $C = U$  και  $X = \Sigma \cdot V^T$ .

Από την καταστατική εξίσωση:

$$X_1^\tau = A \cdot X_0^{\tau-1} + V_1^\tau \quad (A.25)$$

όπου  $V_1^\tau = [v(1) \dots v(\tau)] \in \mathfrak{R}^{n \times \tau}$ , θα υπολογίσουμε την μήτρα  $A$ , έτσι ώστε:

$$\min \|V_1^\tau\|_F \rightarrow \min \|X_1^\tau - A \cdot X_0^{\tau-1}\|_F \quad (A.26)$$

Σύμφωνα με την αρχή της ορθογωνιότητας το σφάλμα εκτίμησης  $X_1^\tau - A \cdot X_0^{\tau-1}$  είναι ορθογώνιο με το δεδομένο  $X_0^{\tau-1}$ . Έτσι έχουμε:



$$(X_1^r - A \cdot X_0^{r-1}) \cdot X_0^{r-1} = 0 \Rightarrow X_1^r - A \cdot X_0^{r-1} = 0 \Rightarrow X_1^r = A \cdot X_0^{r-1} \quad (A.27)$$

Χρησιμοποιώντας τους πίνακες  $D_1 = \begin{bmatrix} 0 & 0 \\ I & 0 \end{bmatrix} \in \mathfrak{R}^{r \times r}$ ,  $D_2 = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} \in \mathfrak{R}^{r \times r}$  και το  $X$  που βρήκαμε παραπάνω τα  $X_1^r, X_0^{r-1}$  γίνονται:

$$X_1^r = \Sigma \cdot V^T \cdot D_1 \text{ και } X_0^{r-1} = \Sigma \cdot V^T \cdot D_2 \quad (A.28)$$

Έτσι η σχέση (A.27) γίνεται:

$$\Sigma \cdot V^T \cdot D_1 = A \cdot \Sigma \cdot V^T \cdot D_2 \quad (A.29)$$

Πολλαπλασιάζοντας με τον ανάστροφο από δεξιά παίρνουμε:

$$\begin{aligned} \Sigma \cdot V^T \cdot D_1 (\Sigma \cdot V^T \cdot D_2)^T &= A \cdot \Sigma \cdot V^T \cdot D_2 (\Sigma \cdot V^T \cdot D_2)^T \Rightarrow \\ \Sigma \cdot V^T \cdot D_1 \cdot D_2^T \cdot V \cdot \Sigma &= A \cdot \Sigma \cdot V^T \cdot D_2 \cdot D_2^T \cdot V \cdot \Sigma \Rightarrow \\ \Sigma \cdot V^T \cdot D_1 \cdot V \cdot \Sigma &= A \cdot \Sigma \cdot V^T \cdot D_2 \cdot V \cdot \Sigma \end{aligned} \quad (A.30)$$

Πολλαπλασιάζοντας από δεξιά με τον αντίστροφο του  $\Sigma \cdot V^T \cdot D_2 \cdot V \cdot \Sigma$  η παραπάνω σχέση γίνεται:

$$\begin{aligned} A &= \Sigma \cdot V^T \cdot D_1 \cdot V \cdot \Sigma \cdot (\Sigma \cdot V^T \cdot D_2 \cdot V \cdot \Sigma)^{-1} \Rightarrow \\ A &= \Sigma \cdot V^T \cdot D_1 \cdot V \cdot \Sigma \cdot (\Sigma^{-1} \cdot (V^T \cdot D_2 \cdot V)^{-1} \cdot \Sigma^{-1}) \Rightarrow \\ A &= \Sigma \cdot V^T \cdot D_1 \cdot V \cdot (V^T \cdot D_2 \cdot V)^{-1} \cdot \Sigma^{-1} \end{aligned} \quad (A.31)$$

Για τον υπολογισμό της μήτρας συνδιασποράς  $Q$  χρησιμοποιούμε τον τύπο:

$$Q = \frac{1}{\tau - 1} \sum_{n=1}^{\tau-1} (x(n+1) - A \cdot x(n)) \cdot (x(n+1) - A \cdot x(n))^T \quad (A.32)$$

Αν χρησιμοποιήσουμε την εξίσωση κατάστασης του μοντέλου (1) θα πρέπει να υπολογίσουμε το  $B$ . Η διαδικασία που ακολουθούμε είναι η κάτωθι:

$$\begin{aligned} (X_1^r - A \cdot X_0^{r-1}) &= B \cdot u(n) \Rightarrow \\ (X_1^r - A \cdot X_0^{r-1}) \cdot (X_1^r - A \cdot X_0^{r-1})^T &= B \cdot u(n) \cdot (B \cdot u(n))^T \Rightarrow \\ (X_1^r - A \cdot X_0^{r-1}) \cdot (X_1^r - A \cdot X_0^{r-1})^T &= B \cdot u(n) \cdot u(n)^T \cdot B^T \Rightarrow \\ (X_1^r - A \cdot X_0^{r-1}) \cdot (X_1^r - A \cdot X_0^{r-1})^T &= \tau \cdot B \cdot \left( \frac{1}{\tau} \cdot u(n) \cdot u(n)^T \right) \cdot B^T \Rightarrow \\ (X_1^r - A \cdot X_0^{r-1}) \cdot (X_1^r - A \cdot X_0^{r-1})^T &= \tau \cdot B \cdot B^T \\ (X_1^r - A \cdot X_0^{r-1}) \cdot (X_1^r - A \cdot X_0^{r-1})^T &= \sqrt{\tau} \cdot B \cdot \sqrt{\tau} \cdot B^T \end{aligned} \quad (A.33)$$

Παίρνοντας το svd του πίνακα  $(X_1^T - A \cdot X_0^{T-1}) = U \cdot \Sigma \cdot V^T$

$$\begin{aligned} (X_1^T - A \cdot X_0^{T-1}) &= U \cdot \Sigma \cdot V^T \Rightarrow \\ (X_1^T - A \cdot X_0^{T-1}) \cdot (X_1^T - A \cdot X_0^{T-1})^T &= U \cdot \Sigma \cdot V^T \cdot V \cdot \Sigma \cdot U^T \Rightarrow \\ (X_1^T - A \cdot X_0^{T-1}) \cdot (X_1^T - A \cdot X_0^{T-1})^T &= U \cdot \Sigma \cdot \Sigma \cdot U^T \Rightarrow \\ (X_1^T - A \cdot X_0^{T-1}) \cdot (X_1^T - A \cdot X_0^{T-1})^T &= U \cdot \Sigma \cdot (U \cdot \Sigma)^T \end{aligned} \quad (A.34)$$

Από τις δύο τελευταίες σχέσεις έχουμε ότι:

$$\sqrt{\tau} \cdot B \cdot \sqrt{\tau} \cdot B^T = U \cdot \Sigma \cdot (U \cdot \Sigma)^T \quad (A.35)$$

Έτσι μπορούμε να επιλέξουμε  $\sqrt{\tau} \cdot B = U \cdot \Sigma \Rightarrow B = \frac{1}{\sqrt{\tau}} \cdot U \cdot \Sigma$ .

## A.9 Παρουσίαση Αποτελεσμάτων

Στο παρόν εδάφιο θα παρουσιάσουμε τα αποτελέσματα της προσομοίωσης του μαθηματικού μοντέλου που αναλύσαμε παραπάνω. Συγκεκριμένα χρησιμοποιήσαμε δύο διαφορετικά βίντεο και πειραματιστήκαμε αλλάζοντας τις τιμές κάποιων παραμέτρων. Για την ακρίβεια αλλάζουμε τον αριθμό των καρέ του αυθεντικού βίντεο που χρησιμοποιεί το μοντέλο για να «μάθει» την κίνηση, αλλάζουμε την διάσταση του διανύσματος κατάστασης και τέλος τροποποιούμε το πλάτος των πόλων του συστήματος έτσι ώστε να αποκτήσει κάποια επιθυμητά χαρακτηριστικά.

Το πρώτο βίντεο έχει ανάλυση  $143 \times 180$  και απεικονίζει μία φωτιά, ενώ το δεύτερο έχει ανάλυση  $323 \times 224$  και απεικονίζει ένα σιντριβάνι. Τα συγκεκριμένα βίντεο βρίσκονται στην βάση δεδομένων του MIT.

### A.9α Βίντεο 1

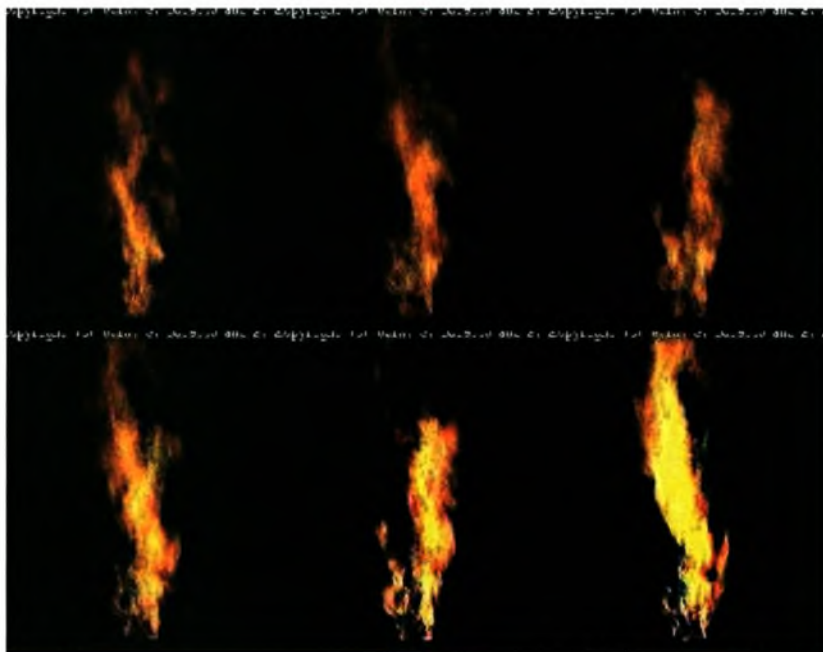
Ακολούθως παρουσιάζουμε κάποια καρέ από την αυθεντική ακολουθία.





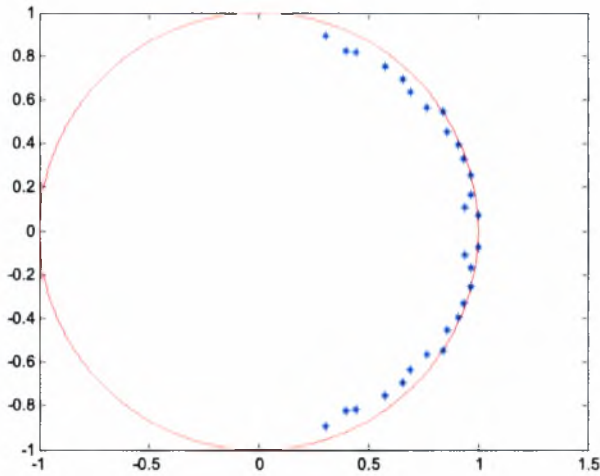
Εικόνα Α.1. Οι παραπάνω εικόνες αποτελούν από αριστερά προς τα δεξιά και από πάνω προς τα κάτω το  $1^\circ$ , το  $10^\circ$ , το  $30^\circ$ , το  $60^\circ$ , το  $70^\circ$  και το  $90^\circ$  καρέ του αυθεντικού βίντεο.

Παρακάτω παρουσιάζουμε τα αποτελέσματα της προσομοίωσης για διάσταση διανύσματος κατάστασης ίση με 30 και μέγεθος συνόλου εκμάθησης ίσο με 90.

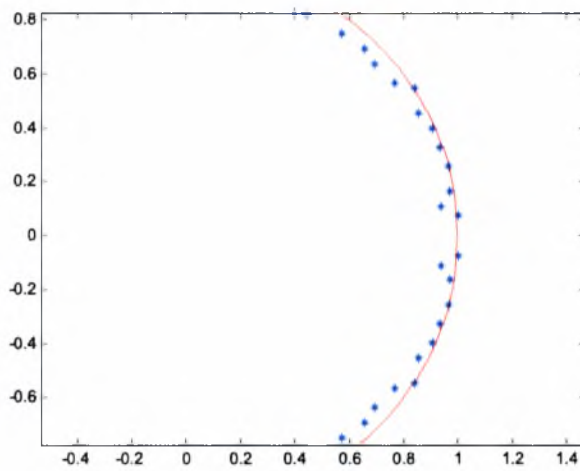


Εικόνα Α.2α. Οι παραπάνω εικόνες αποτελούν το  $10^\circ$ ,  $50^\circ$ ,  $150^\circ$ ,  $300^\circ$ ,  $400^\circ$ ,  $550^\circ$  καρέ αντίστοιχα που συνθέσα για μέγεθος διανύσματος κατάστασης ίσο με 30.

Παρατηρούμε ότι η οπτική ποιότητα των καρέ μειώνεται όταν συνθέτω βίντεο μεγάλης διάρκειας. Ο λόγος για τον οποίο αποκλίνει το σύστημά μας είναι επειδή έχει πόλους εκτός του μοναδιαίου κύκλου. Ακολουθεί το γράφημα που απεικονίζει τους πόλους του συστήματος.



Εικόνα Α.2β. Πόλοι συστήματος για μέγεθος διανύσματος κατάστασης ίσο με 30.



Εικόνα Α.2γ. Πόλοι συστήματος για μέγεθος διανύσματος κατάστασης ίσο με 30 κάνοντας zoom.

Παρατηρούμε ότι οι πόλοι που βρίσκονται εκτός του μοναδιαίου κύκλου έχουν μέτρο λίγο μεγαλύτερο από τη μονάδα και έτσι αν τους μεταφέρουμε εντός μοναδιαίου κύκλου μειώνοντας το μέτρο τους και κάνοντάς το 0.999 και διατηρώντας τη φάση ίδια η μήτρα του συστήματος τροποποιείται λίγο και το σύστημά αποκτά την επιθυμητή ιδιότητα της ευστάθειας.

Συγκεκριμένα εφαρμόζουμε αποδόμηση ιδιοτιμών στη μήτρα  $A$  του συστήματος:

$$A = T \cdot \Lambda \cdot T^{-1} \quad (A.36)$$

όπου οι στήλες του  $T$  είναι τα ιδιοδιανύσματα του πίνακα  $A$  και  $\Lambda$  διαγώνιος πίνακας που περιέχει τις ιδιοτιμές της μήτρας του συστήματος. Όπως έχουμε αναφέρει τα στοιχεία του πίνακα  $\Lambda$  αποτελούν τους πόλους του συστήματος. Έτσι υπολογίζουμε το μέτρο των διαγώνιων στοιχείων του και το τροποποιούμε όταν είναι

μεγαλύτερο της μονάδας. Δηλαδή αν  $z$  μία ιδιοτιμή του πίνακα  $A$  και  $|z| > 1$ , τότε  $z = 0.999 \cdot \angle z$ , όπου  $\angle z$  η φάση του  $z$ . Αφού ο πίνακας  $\Lambda$  έχει αλλάξει και έχει γίνει  $\Lambda'$  υπολογίζω ξανά τον πίνακα  $A$  ως εξής:

$$A' = T \cdot \Lambda' \cdot T^{-1}$$

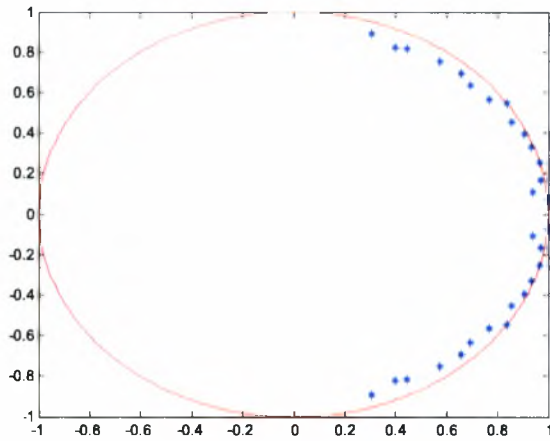
Ο πίνακας  $A'$  αποτελεί πλέον τη νέα μήτρα του συστήματος προσδίδοντάς του την ιδιότητα της ευστάθειας.

Παρακάτω παρουσιάζουμε τα αποτελέσματα της προσομοίωσης χρησιμοποιώντας τη νέα μήτρα συστήματος που την υπολογίσαμε με τον τρόπο που μόλις περιγράψαμε.



Εικόνα Α.3α. Οι παραπάνω εικόνες αποτελούν το  $10^\circ$ ,  $50^\circ$ ,  $150^\circ$ ,  $300^\circ$ ,  $400^\circ$ ,  $550^\circ$  καρέ αντίστοιχα που συνθέσα για μέγεθος διανύσματος κατάστασης ίσο με 30.

Παρατηρούμε ότι το οπτικό αποτέλεσμα είναι καλύτερο σε σύγκριση με τα καρέ που παρουσιάσαμε προηγουμένως κάτι που οφείλεται βέβαια στο ότι το σύστημά μας είναι πλέον ευσταθές. Οι πόλοι του συστήματος είναι οι εξής:



Εικόνα Α.3β. Πόλοι συστήματος για μέγεθος διανύσματος κατάσταση ίσο με 30.

### Α.9α.1 Παρουσίαση αποτελεσμάτων αλλάζοντας το μέγεθος του διανύσματος κατάσταση

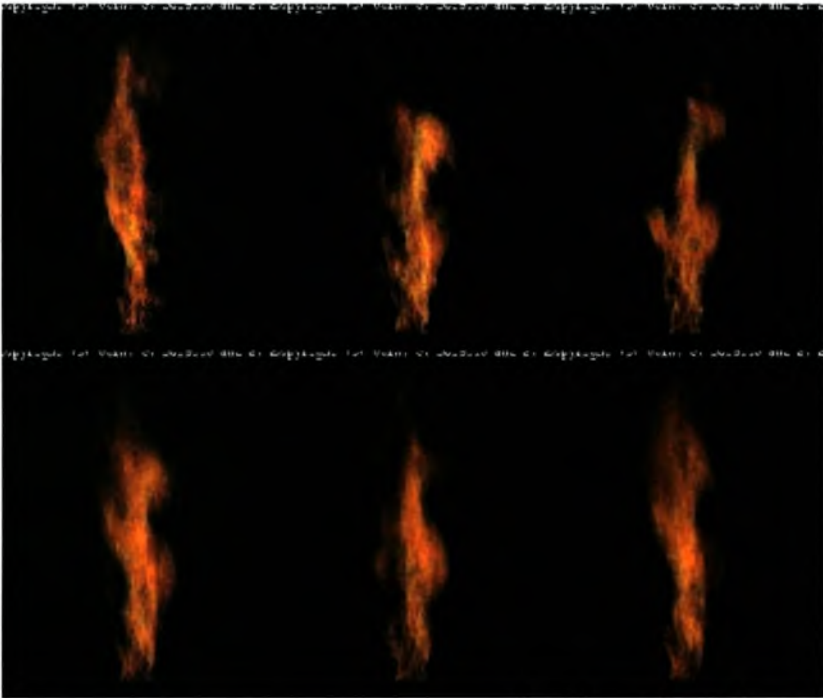
Στην παρούσα παράγραφο θα παρουσιάσουμε τα αποτελέσματα της προσομοίωσης χρησιμοποιώντας τρία διαφορετικά μεγέθη διανύσματος κατάσταση. Συγκεκριμένα τα μεγέθη που θα χρησιμοποιήσουμε είναι 40, 60 και 80. Ως μέτρο σύγκρισης θα χρησιμοποιήσουμε την οπτική ποιότητα των βίντεο που συνθέσαμε. Θα πρέπει στο σημείο αυτό να αναφέρουμε ότι η ποιοτική διαφορά των καρτέ που συνθέσαμε για διαφορετικό μήκος διανύσματος κατάσταση θα γινόταν περισσότερο ευδιάκριτη αν παρακολουθούσαμε ολόκληρο το βίντεο και όχι απλά κάποια από τα καρτέ του. Τέλος το σύνολο εκμάθησης αποτελείται από 100 εικόνες και οι πόλοι του συστήματος έχουν ελεγχθεί με τον προαναφερθέντα τρόπο. Οι εικόνες που πήραμε για μέγεθος διανύσματος κατάσταση ίσο με 40 είναι:





Εικόνα Α.4. Οι παραπάνω εικόνες αποτελούν το  $10^\circ$ ,  $50^\circ$ ,  $150^\circ$ ,  $300^\circ$ ,  $400^\circ$ ,  $550^\circ$  καρέ αντίστοιχα που συνθέσα για μέγεθος διανύσματος κατάσταση ίσο με 40.

Οι εικόνες που πήραμε για μέγεθος διανύσματος κατάσταση ίσο με 60 είναι:



Εικόνα Α.5. Οι παραπάνω εικόνες αποτελούν το  $10^\circ$ ,  $50^\circ$ ,  $150^\circ$ ,  $300^\circ$ ,  $400^\circ$ ,  $550^\circ$  καρέ αντίστοιχα που συνθέσα για μέγεθος διανύσματος κατάσταση ίσο με 60.

Οι εικόνες που πήραμε για μέγεθος διανύσματος κατάσταση ίσο με 80 είναι:







Εικόνα Α.6. Οι παραπάνω εικόνες αποτελούν το  $10^\circ$ ,  $50^\circ$ ,  $150^\circ$ ,  $300^\circ$ ,  $400^\circ$ ,  $550^\circ$  καρέ αντίστοιχα που συνθέσα για μέγεθος διανύσματος κατάστασης ίσο με 60.

Παρατηρούμε ότι οπτική ποιότητα του βίντεο είναι καλύτερη στην περίπτωση που χρησιμοποιούμε μεγαλύτερο μήκος διανύσματος κατάστασης. Παρακολουθώντας τα βίντεο είδαμε ότι όταν το μήκος αυξάνεται το σύστημα «συλλαμβάνει» περισσότερες λεπτομέρειες από το βίντεο μάθησης με αποτέλεσμα οι εικόνες να έχουν υψηλότερη ποιότητα. Η διαφορά αυτή είναι περισσότερο εμφανής όταν τροποποιώ το μήκος από 40 σε 60 από όταν το αλλάζω από 60 σε 80.

### Α.9α.2 Παρουσίαση χρονικής απόδοσης της προσομοίωσης

Στην προηγούμενη παράγραφο αναφέραμε ότι καθώς αυξάνουμε το μέγεθος του διανύσματος κατάστασης η ποιότητα του βίντεο βελτιώνεται. Ωστόσο όταν θέλουμε να συνθέσουμε καρέ σε πραγματικό χρόνο μας ενδιαφέρει να δημιουργήσουμε νέες εικόνες με όσο το δυνατόν μεγαλύτερο ρυθμό κάτι το οποίο απαιτεί σχετικά μικρό μήκος διανύσματος κατάστασης. Αυτό βέβαια δεν είναι απόλυτο καθώς ο ρυθμός με τον οποίο μπορεί να γίνει σύνθεση νέων καρέ είναι συνάρτηση της υπολογιστικής ισχύος του εκάστοτε υπολογιστικού συστήματος. Έτσι θα πρέπει να βρούμε μια μέση λύση η οποία να συνδυάζει το καλό οπτικό αποτέλεσμα με έναν αρκετά μεγάλο ρυθμό.

Οι προσομοιώσεις έγιναν σε σύστημα με υπολογιστική ισχύ 3.2 GH (Pentium IV) και μνήμη RAM 512 MB.

Ακολουθώς θα παρουσιάσουμε το χρόνο που κάνει το σύστημα να συνθέσει 200 καρέ και το χρόνο που απαιτείται για να μάθει 100 εικόνες. Αν υποθέσουμε ότι ο ιδανικός ρυθμός σύνθεσης καρέ είναι 25, τότε το σύστημα θα πρέπει να συνθέτει  $\frac{200}{25} = 8$  καρέ το δευτερόλεπτο.

#### Α.9α.2α Χρονική απόδοση του συστήματος κατά το στάδιο της σύνθεσης

Ακολουθεί ο πίνακας που παρουσιάζει το μήκος του διανύσματος κατάστασης, το χρόνο που απαιτείται για την σύνθεση 200 εικόνων και το ρυθμό σύνθεσης καρέ που

ορίζεται ως  $r = \frac{200}{\text{χρονος} \cdot \text{συνθεσης}}$ .



Μήκος διανύσματος κατάστασης	Χρόνος σύνθεσης	Ρυθμός σύνθεσης
15	4.62	43.29
20	5.4	37.03
25	6.15	32.52
30	6.79	29.45
40	8.29	24.12
50	9.59	20.85
60	11.04	18.11

Πίνακας Α.1.

Παρατηρούμε όπως ήταν βέβαια αναμενόμενο ότι ο χρόνος σύνθεσης είναι ανάλογος με το μήκος διανύσματος.

### Α.9α.2β Χρονική απόδοση του συστήματος κατά το στάδιο της εκμάθησης

Ακολουθεί ο πίνακας στον οποίο παρουσιάζεται το μέγεθος του συνόλου εκμάθησης και ο χρόνος ο οποίος απαιτείται για να μάθει το μοντέλο την κίνηση. Στους χρόνους που παρουσιάζουμε δεν συμπεριλαμβάνεται ο χρόνος που απαιτείται για να διαβαστούν οι εικόνες, αλλά υποθέτουμε ότι όλα τα καρέ βρίσκονται σε ένα πίνακα στην RGB μορφή τους. Επίσης η χρονική απόδοση του συστήματος υπολογίζεται για τρία διαφορετικά μήκη διανύσματος κατάστασης. Συγκεκριμένα τα μήκη που χρησιμοποιούμε είναι ίσα με 40, 60 και 80.

Πλήθος καρέ	Χρόνος εκμάθησης		
	40	60	80
100	8.36	8.44	8.5
95	7.75	7.83	7.86
90	6.95	7.03	7.05
85	6.47	6.28	6.34
80	5.52	5.59	5.65

Πίνακας Α.2.

### Α.9β Βίντεο 2

Στο παρόν εδάφιο θα παρουσιάσουμε τα αποτελέσματα που πήραμε χρησιμοποιώντας το δεύτερο βίντεο που απεικονίζει ένα σιντριβάνι. Παρακάτω παραθέτουμε κάποια καρέ από την αυθεντική ακολουθία.



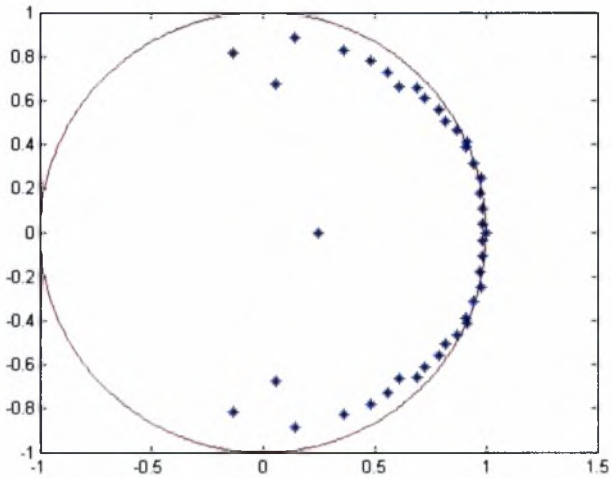
Εικόνα Α.7. Οι παραπάνω εικόνες αποτελούν από αριστερά προς τα δεξιά και από πάνω προς τα κάτω το  $1^\circ$  ,  $30^\circ$  ,  $60^\circ$  ,  $90^\circ$  καρέ του αυθεντικού βίντεο.

Χρησιμοποιώντας μέγεθος διανύσματος κατάστασης ίσο με 40 και χωρίς να επέμβουμε στους πόλους του συστήματος συνθέτουμε βίντεο και παίρνουμε τα παρακάτω αποτελέσματα.



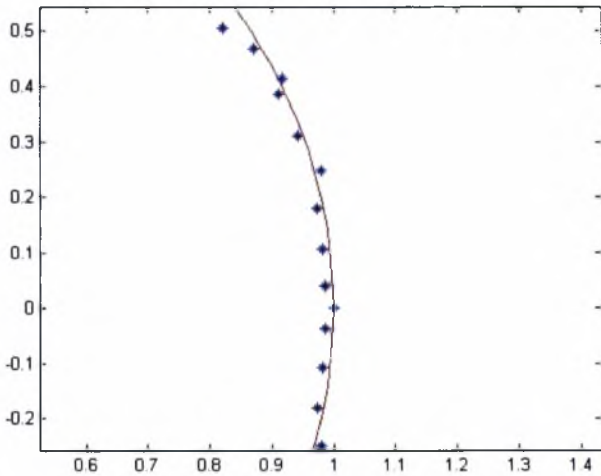
Εικόνα Α.8α. Οι παραπάνω εικόνες αποτελούν το  $50^\circ$  ,  $200^\circ$  ,  $400^\circ$  ,  $600^\circ$  καρέ αντίστοιχα που συνθέσα για μέγεθος διανύσματος κατάστασης ίσο με 40.

Παρατηρούμε ότι η μείωση της οπτικής ποιότητας των εικόνων γίνεται περισσότερο αισθητή όταν συνθέτουμε βίντεο μεγάλης διάρκειας. Η αιτία αυτού του φαινομένου, όπως στην περίπτωση της φωτιάς, εντοπίζεται στους πόλους του συστήματος οι οποίοι απεικονίζονται στο παρακάτω γράφημα.



Εικόνα Α.8β. Πόλοι του συστήματος

Παρακάτω παρουσιάζουμε πάλι τους πόλους του συστήματος κάνοντας zoom στην περιοχή στην οποία οι πόλοι βρίσκονται εκτός του μοναδιαίου κύκλου.



Εικόνα Α.8γ. Πόλοι του συστήματος

Παρατηρούμε ότι τέσσερις πόλοι βρίσκονται εκτός μοναδιαίου κύκλου με αποτέλεσμα το σύστημά μας να είναι ασταθές και έτσι το οπτικό αποτέλεσμα των εικόνων να χειροτερεύει με το πέρασμα των καρέ.

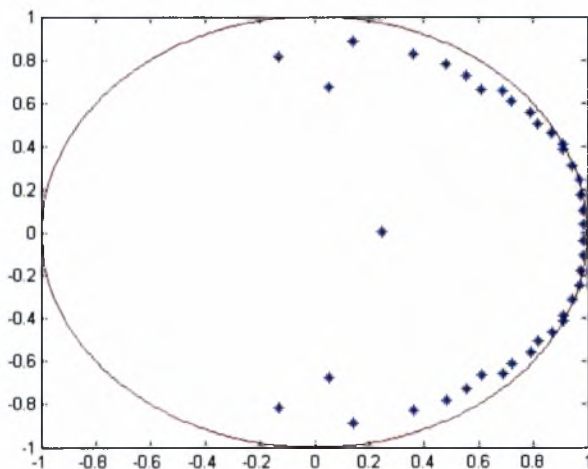


Ακολουθώντας την ίδια διαδικασία όπως στην περίπτωση της φωτιάς, δηλαδή αλλάζοντας το μέτρο των ιδιοτιμών τις μήτρας του συστήματος που είναι μεγαλύτερο από μονάδα σε 1, παίρνουμε τα παρακάτω αποτελέσματα.



Εικόνα Α.9α. Οι παραπάνω εικόνες αποτελούν το  $50^\circ$ ,  $200^\circ$ ,  $400^\circ$ ,  $600^\circ$  καρέ αντίστοιχα που συνθέσα για μέγεθος διανύσματος κατάστασης ίσο με 40.

Παρατηρούμε ότι η οπτική ποιότητα των καρέ δεν μειώνεται ακόμη και όταν συνθέτουμε βίντεο μεγάλης διάρκειας. Ακολουθεί το γράφημα με τους τροποποιημένους πόλους.



Εικόνα Α.9β. Οι τροποποιημένοι πόλοι του συστήματος

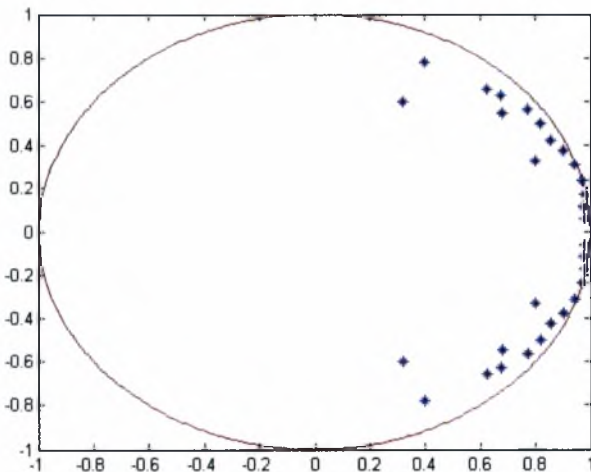
### A.9β.1 Παρουσίαση αποτελεσμάτων αλλάζοντας το μέγεθος του διανύσματος κατάστασης

Στην παράγραφο αυτή θα παρουσιάσουμε τις εικόνες που παίρνουμε αν αλλάξουμε το μέγεθος του διανύσματος κατάστασης. Συγκεκριμένα θα χρησιμοποιήσουμε δύο διαφορετικά μήκη διανύσματος κατάστασης (30, 60) και θα δούμε τι επίδραση θα έχει η αλλαγή του μήκους στα καρέ που συνθέτουμε.

Ακολουθούν τα καρέ που συνθέσαμε για μήκος διανύσματος κατάστασης ίσο με 30.



Εικόνα A.10α. Οι παραπάνω εικόνες αποτελούν το  $300^\circ$ ,  $301^\circ$ ,  $302^\circ$ ,  $303^\circ$  καρέ αντίστοιχα που συνθέσαμε για μέγεθος διανύσματος κατάστασης ίσο με 30.

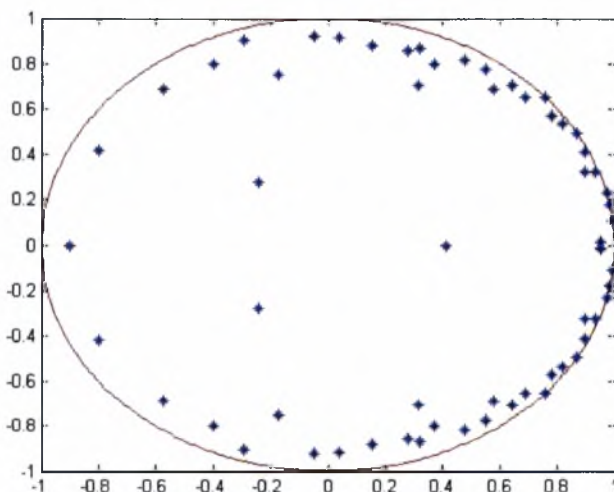


Εικόνα A.10β. Οι πόλοι του συστήματος για μέγεθος διανύσματος κατάστασης ίσο με 30.

Τα καρέ που πήραμε για μέγεθος διανύσματος κατάσταση ίσο με 60 είναι:



Εικόνα Α.11α. Οι παραπάνω εικόνες αποτελούν το  $300^\circ$ ,  $301^\circ$ ,  $302^\circ$ ,  $303^\circ$  καρέ αντίστοιχα που συνθέσαμε για μέγεθος διανύσματος κατάσταση ίσο με 30.



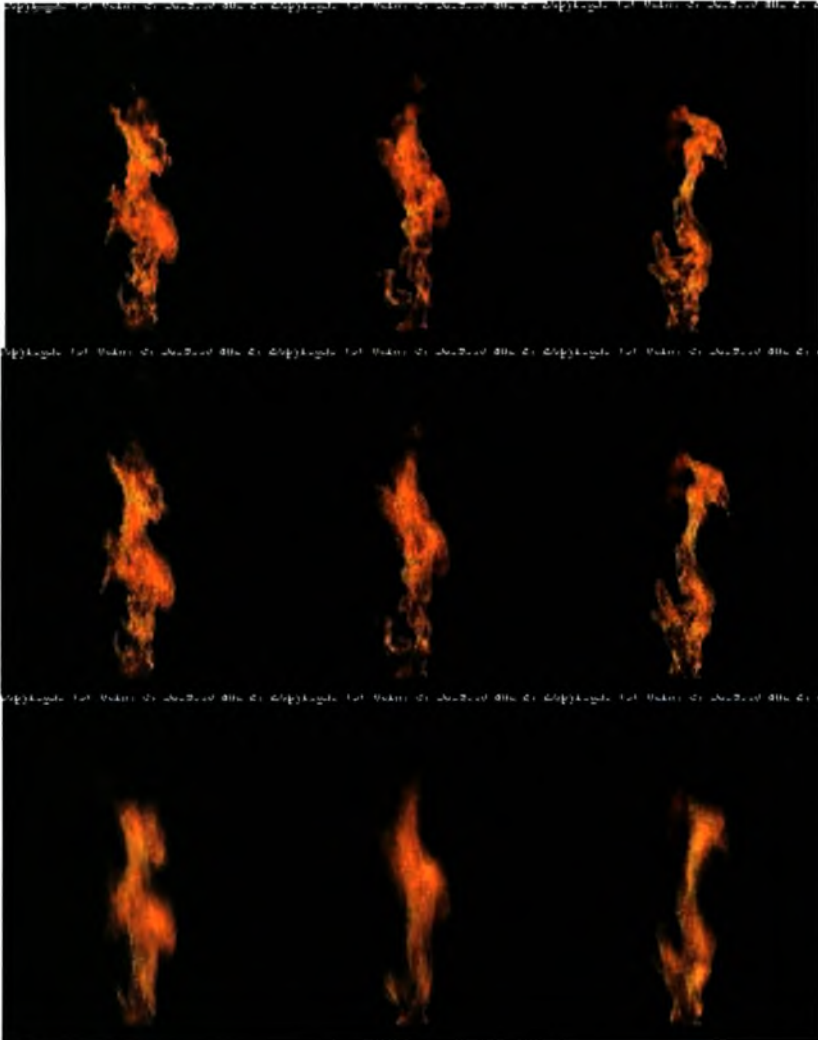
Εικόνα Α.11β. Οι πόλοι του συστήματος για μέγεθος διανύσματος κατάσταση ίσο με 60.

Παρατηρούμε ότι στην περίπτωση που χρησιμοποιούμε μέγεθος διανύσματος κατάσταση ίσο με 60 το βίντεο που συνθέτουμε είναι καλύτερο από την άποψη ότι το σιντριβάνι παρουσιάζει μεγαλύτερη ποικιλία κινήσεων. Αυτό οφείλεται στο γεγονός ότι το σύστημα «συγκρατεί» περισσότερες λεπτομέρειες όταν χρησιμοποιούμε μήκος ίσο με 60 παρά όταν το μήκος είναι ίσο με 30.



## A.10 Συμπύση

Το μοντέλο των Gianfranco Doretto, Alessandro Chiuso, Ying N.Wu και Stefano Soatto μπορεί να χρησιμοποιηθεί για την συμπύση των εικόνων. Συγκεκριμένα για την αποθήκευση  $\tau$  εικόνων απαιτείται η αποθήκευση  $\tau \cdot m$  (A.37) αριθμών, όπου  $m$  το πλήθος των εικονοστοιχείων της κάθε εικόνας, ενώ με το μοντέλο που περιγράψαμε χρειάζεται η αποθήκευση των εξής παραμέτρων: η μήτρα του συστήματος  $A$  ( $n^2$  αριθμοί), ο πίνακας  $B$  ( $n \cdot n$ , αριθμοί), η ακολουθία εισόδου  $u(n)$  ( $n_v \cdot \tau$  αριθμοί) και η μήτρα εξόδου  $C$  ( $m \cdot n$  αριθμοί). Συνολικά απαιτείται η αποθήκευση  $n^2 + n \cdot n_v + n_v \cdot \tau + m \cdot n$  (A.38) αριθμών. Από τις σχέσεις (A.37) και (A.38) γίνεται αντιληπτό ότι όσο μεγαλύτερο είναι το πλήθος των εικόνων τόσο μεγαλύτερος θα είναι και ο λόγος συμπύσης. Ακολουθούν οι εικόνες που παίρνουμε χρησιμοποιώντας μήκος διανύσματος κατάστασης ίσο με 50 και 15. Στην πρώτη σειρά παραθέτουμε τις εικόνες της αυθεντικής ακολουθίας, στην δεύτερη σειρά τα καρέ που πήραμε χρησιμοποιώντας μήκος διανύσματος κατάστασης ίσο με 50 και στη τρίτη σειρά τα καρέ που πήραμε χρησιμοποιώντας μήκος διανύσματος κατάστασης ίσο με 15.

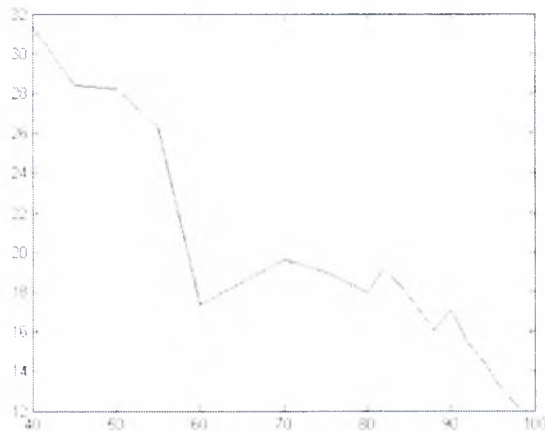


Εικόνα A.12

Παρατηρούμε, όπως είναι λογικό, ότι η οπτική ποιότητα των καρτέ μειώνεται χρησιμοποιώντας μικρότερο μήκος διανύσματος κατάστασης. Αυτό είναι το κόστος για την επίτευξη καλύτερου λόγου συμπίεσης. Χρησιμοποιώντας μήκος διανύσματος κατάστασης ίσο με 50 ο λόγος συμπίεσης είναι περίπου 2, ενώ αν το διάνυσμα κατάστασης έχει μήκος ίσο με 15 ο λόγος συμπίεσης είναι περίπου 6.6.

### A.11 Αλλαγή του μεγέθους του συνόλου εκμάθησης και του πλήθους των διανυσμάτων βάσης

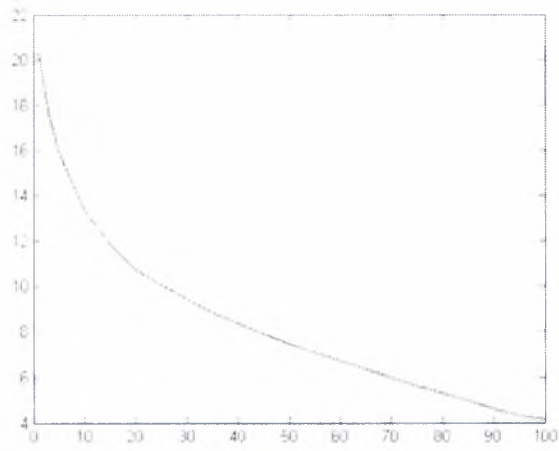
Όπως ήδη έχουμε αναφέρει το σύστημα χρησιμοποιεί έναν αριθμό από καρτέ για τον υπολογισμό των παραμέτρων. Χρησιμοποιώντας διαφορετικό πλήθος εικόνων το αποτέλεσμα αλλάζει. Συγκεκριμένα όταν χρησιμοποιούμε περισσότερες εικόνες το σφάλμα πρόβλεψης μειώνεται. Η διαδικασία που ακολουθούμε είναι η εξής: χρησιμοποιούμε  $x$  καρτέ από την αυθεντική ακολουθία εικόνων και συγκρίνουμε το πρώτο καρτέ των εικόνων που συνθέσαμε με το  $x+1$  καρτέ της αυθεντικής ακολουθίας. Δηλαδή αφαιρούμε τις RGB τιμές των δύο καρτέ και διαιρούμε με το πλήθος των εικονοστοιχείων. Το γράφημα που παίρνουμε για την περίπτωση της φωτιάς είναι:



Εικόνα A.13 Σφάλμα πρόβλεψης

Στην προηγούμενη παράγραφο αναφέραμε ότι το μοντέλο μπορεί να χρησιμοποιηθεί για την συμπίεση των εικόνων. Το γεγονός ότι το αποτέλεσμα διαφέρει χρησιμοποιώντας διαφορετικό πλήθος διανυσμάτων βάσης μπορεί να φανεί χρησιμοποιώντας το σφάλμα συμπίεσης. Συγκεκριμένα το σφάλμα μειώνεται καθώς αυξάνουμε το πλήθος των διανυσμάτων βάσης. Για τον υπολογισμό του σφάλματος συμπίεσης αφαιρούμε τις RGB τιμές των αυθεντικών εικόνων από τις αντίστοιχες των συμπίεσμένων εικόνων και διαιρούμε με το πλήθος των εικονοστοιχείων επί το πλήθος των εικόνων. Η γραφική παράσταση που παίρνουμε για την περίπτωση της φωτιάς είναι:





Εικόνα Α.14 Σφάλμα συμπίεσης

## ΜΕΡΟΣ Β

### B.1 Εισαγωγή

Όπως αναφέραμε στην αρχή της εργασίας στην γενική περίπτωση το μοντέλο περιγράφεται από τις παρακάτω σχέσεις.

$$\begin{aligned}x(n) &= \sum_{i=1}^k A_i x(n-i) + u(n) \\y(n) &= Cx(n) + w(n)\end{aligned}$$

Στο παρόν εδάφιο θα παρουσιάσουμε τα αποτελέσματα που πήραμε χρησιμοποιώντας τρεις προηγούμενες καταστάσεις. Έτσι οι παραπάνω σχέσεις γίνονται:

$$\begin{aligned}x(n) &= A_1 x(n-1) + A_2 x(n-2) + A_3 x(n-3) + u(n) \\y(n) &= Cx(n)\end{aligned} \quad (B.1)$$

Το μοντέλο (1) μπορεί να ξαναγραφεί ως:

$$\begin{aligned}x(n) &= A \cdot X + u(n) \\y(n) &= Cx(n) + w(n)\end{aligned} \quad (B.2)$$

Όπου  $A = [A_1 \quad A_2 \quad A_3]$  και  $X = \begin{bmatrix} X_{n-1} \\ X_{n-2} \\ X_{n-3} \end{bmatrix}$ .

### B.2 Υπολογισμός των παραμέτρων

Για τον υπολογισμό της παραμέτρου  $C$  και των διανυσμάτων κατάστασης ακολουθούμε ακριβώς την ίδια διαδικασία με την περίπτωση που χρησιμοποιούμε μόνο μία προηγούμενη κατάσταση. Για τον υπολογισμό του πίνακα  $A$  χρησιμοποιούμε τα διανύσματα  $x$  που υπολογίσαμε στην πρώτη φάση και δημιουργούμε ένα πίνακα, έστω  $X'$  ο οποίος είναι ίσος με:

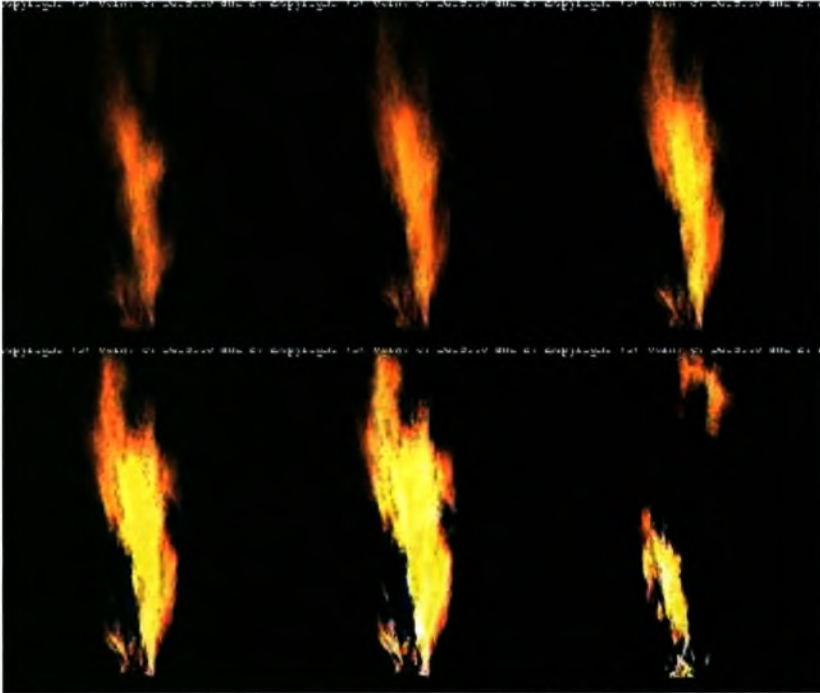
$$X' = \begin{bmatrix} X_1 & X_2 & X_3 & X_4 & \cdots & X_k \\ 0 & X_1 & X_2 & X_3 & \cdots & X_{k-1} \\ 0 & 0 & X_1 & X_2 & \cdots & X_{k-2} \end{bmatrix} \quad (B.3)$$

όπου  $k$  το πλήθος των εικόνων που χρησιμοποιούμε κατά την διαδικασία της μάθησης. Κατόπιν ακολουθούμε την ίδια διαδικασία που χρησιμοποιήσαμε και στο πρώτο μέρος της εργασίας με την διαφορά ότι ο πίνακας  $A$  είναι διάστασης  $m \times 3m$ , όπου  $m$  το μήκος του διανύσματος κατάστασης, και το  $X'$  είναι διάστασης  $3m \times k$ .

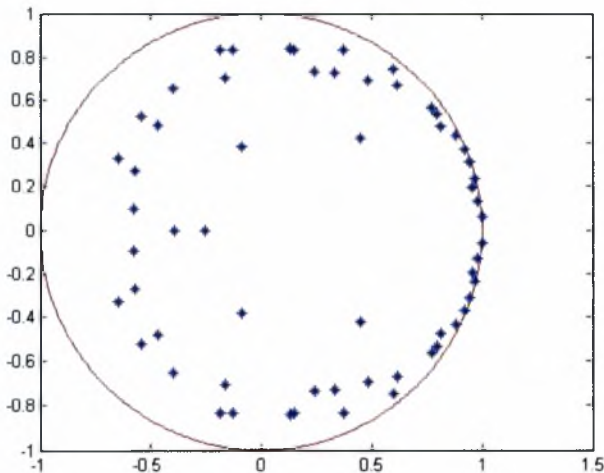
Πρέπει να προσέξουμε όταν πάρουμε τον ψευδοαντίστροφο του  $X'$ , καθώς ο ψευδοαντίστροφος αυτός υπάρχει υπό την προϋπόθεση ότι  $3m > k$ . Για να επιτύχουμε κάτι τέτοιο θα πρέπει είτε να επιλέξουμε μικρό μήκος διανύσματος κατάστασης είτε μεγάλο πλήθος εικόνων.

### B.3 Παρουσίαση αποτελεσμάτων

Εφαρμόζοντας λοιπόν το παραπάνω μοντέλο παίρνουμε τα κάτωθι αποτελέσματα.



Εικόνα B.2α. Οι παραπάνω εικόνες αποτελούν το  $50^\circ$ ,  $150^\circ$ ,  $250^\circ$ ,  $350^\circ$ ,  $450^\circ$ ,  $500^\circ$  καρέ αντίστοιχα που συνθέσαμε για μέγεθος διανύσματος κατάστασης ίσο με 20.



Εικόνα B.2β. Οι πόλοι του συστήματος

Παρατηρούμε ότι το σύστημα αποκλίνει με συνέπεια το οπτικό αποτέλεσμα των εικόνων που συνθέσαμε να μην είναι καλό. Από το γράφημα με τους πόλους του συστήματος βλέπουμε ότι δύο πόλοι βρίσκονται εκτός μοναδιαίου κύκλου τους οποίους πρέπει να «διορθώσουμε» έτσι ώστε να είναι καλή η οπτική ποιότητα των καρτέ.

#### B.4 Τροποποίηση των πόλων του συστήματος

Ο πίνακας που θέλουμε να εκτιμήσουμε έχει διάσταση  $m \times 3m$  και έτσι δεν μπορούμε να κάνουμε ανάλυση ιδιοτιμών. Για το λόγο αυτό δημιουργούμε τον παρακάτω πίνακα:

$$A = \begin{bmatrix} A_1 & A_2 & A_3 \\ I & 0 & 0 \\ 0 & I & 0 \end{bmatrix} \quad (B.4)$$

Έτσι ισχύει ότι:

$$\begin{bmatrix} X_{n+1} \\ X_n \\ X_{n-1} \end{bmatrix} = \begin{bmatrix} A_1 & A_2 & A_3 \\ I & 0 & 0 \\ 0 & I & 0 \end{bmatrix} \begin{bmatrix} X_n \\ X_{n-1} \\ X_{n-2} \end{bmatrix} \quad (B.5)$$

Αν κάνουμε ανάλυση ιδιοτιμών στον πίνακα  $A$ , με  $A = T \cdot \Lambda \cdot T^{-1}$ , τροποποιήσουμε τις ιδιοτιμές των οποίων το μέτρο είναι μεγαλύτερο της μονάδας και τελικά πάρουμε  $A' = T \cdot \Lambda' \cdot T^{-1}$ , όπου  $\Lambda'$  ο πίνακας με τις τροποποιημένες ιδιοτιμές, τότε το σύστημα θα είναι ευσταθές αλλά δεν θα ακολουθεί το αρχικό μοντέλο καθώς θα έχουν τροποποιηθεί οι μοναδιαίοι και οι μηδενικοί υποπίνακες του πίνακα  $A$ . Για τον παραπάνω λόγο πρέπει να βρούμε ένα διαφορετικό τρόπο για να τροποποιήσουμε τις ιδιοτιμές του πίνακα  $A$ .

Η σχέση

$$x(n) = A_1 x(n-1) + A_2 x(n-2) + A_3 x(n-3) + u(n) \quad (B.6)$$

μπορεί να γραφτεί και ως:

$$(\lambda^3 I - \lambda^2 A_1 - \lambda A_2 - A_3)X = 0 \quad (B.7)$$

Η παραπάνω σχέση για να ισούται με μηδέν θα πρέπει να ισχύει:

$$\det(\lambda^3 I - \lambda^2 A_1 - \lambda A_2 - A_3) = 0 \quad (B.8)$$

Για να βρούμε τις ιδιοτιμές του πίνακα  $A$  θα πρέπει να υπολογίσουμε την ορίζουσα :

$$\det \left( \mu \begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix} - \begin{bmatrix} A_1 & A_2 & A_3 \\ I & 0 & 0 \\ 0 & I & 0 \end{bmatrix} \right) = 0 \quad (B.9)$$

Έτσι λοιπόν έχουμε:

$$\det \begin{pmatrix} \mu I - A_1 & -A_2 & -A_3 \\ -I & \mu I & 0 \\ 0 & -I & \mu I \end{pmatrix} = 0 \Rightarrow \det((\mu I - A_1)\mu^2 I + I(-A_2\mu - A_3)) = 0 \Rightarrow$$

$$\det(\mu^3 I - \mu^2 A_1 - \mu A_2 - A_3) = 0 \quad (B.10)$$

Από τα παραπάνω συμπεραίνουμε ότι οι ιδιοτιμές του πίνακα  $A$  και τα  $\lambda$  που λύνουν την εξίσωση (2) είναι ίδια.

Επίσης ξέρουμε ότι ισχύει:

$$\begin{bmatrix} A_1 & A_2 & A_3 \\ I & 0 & 0 \\ 0 & I & 0 \end{bmatrix} \begin{bmatrix} Y_{1i} \\ Y_{2i} \\ Y_{3i} \end{bmatrix} = \lambda_i \begin{bmatrix} Y_{1i} \\ Y_{2i} \\ Y_{3i} \end{bmatrix} \quad (B.11)$$

όπου  $\lambda_i$  οι ιδιοτιμές και  $\begin{bmatrix} Y_{1i} \\ Y_{2i} \\ Y_{3i} \end{bmatrix}$  τα ιδιοδιανύσματα του πίνακα  $A$ , με  $i = 1 \dots 3m$ .

Κάνοντας τις πράξεις στην σχέση (3) παίρνουμε:

$$A_1 Y_{1i} + A_2 Y_{2i} + A_3 Y_{3i} = \lambda_i Y_{1i} \quad (B.12)$$

$$Y_{1i} = \lambda_i Y_{2i} \quad (B.13)$$

$$Y_{2i} = \lambda_i Y_{3i} \quad (B.14)$$

Αντικαθιστώντας την σχέση (B.14) στην (B.13) παίρνουμε:

$$Y_{1i} = \lambda_i^2 Y_{3i} \quad (B.15)$$

Αντικαθιστώντας την σχέση (B.15) και την σχέση (B.14) στην (B.12) παίρνουμε:

$$A_1 \lambda_i^2 Y_{3i} + A_2 \lambda_i Y_{3i} + A_3 Y_{3i} = \lambda_i^3 Y_{3i} \quad (B.16)$$

Βγάζοντας κοινό παράγοντα το  $Y_{3i}$  παίρνουμε:

$$(\lambda_i^3 I - A_1 \lambda_i^2 - A_2 \lambda_i - A_3) Y_{3i} = 0 \quad (B.17)$$

Από την τελευταία σχέση συμπεραίνουμε ότι το τρίτο μέρος των ιδιοδιανυσμάτων του πίνακα  $A$  ταυτίζεται με το διάνυσμα  $X$  της σχέσης (2). Έτσι για να υπολογίσουμε τους πίνακες  $A_1, A_2, A_3$  υπολογίζουμε τις ιδιοτιμές και τα ιδιοδιανύσματα του πίνακα  $A$ , τροποποιούμε τους πόλους που προκαλούν την αστάθεια και παίρνουμε το τρίτο μέρος των ιδιοδιανυσμάτων. Δηλαδή θα πρέπει να επιλύσουμε το παρακάτω σύστημα.

$$\begin{bmatrix} \lambda_1^3 Y_{31} & \lambda_2^3 Y_{32} & \cdots & \lambda_{3m}^3 Y_{3,3m} \end{bmatrix} - \begin{bmatrix} A_1 & A_2 & A_3 \end{bmatrix} \begin{bmatrix} \lambda_1^2 Y_{31} & \lambda_2^2 Y_{32} & \cdots & \lambda_{3m}^2 Y_{3,3m} \\ \lambda_1 Y_{31} & \lambda_2 Y_{32} & \cdots & \lambda_{3m} Y_{3,3m} \\ Y_{31} & Y_{32} & \cdots & Y_{3,3m} \end{bmatrix} = 0 \quad (B.18)$$

Κατόπιν πολλαπλασιάζουμε από δεξιά με τον ψευδοαντίστροφο του πίνακα

$$\begin{bmatrix} \lambda_1^2 Y_{31} & \lambda_2^2 Y_{32} & \cdots & \lambda_{3m}^2 Y_{3,3m} \\ \lambda_1 Y_{31} & \lambda_2 Y_{32} & \cdots & \lambda_{3m} Y_{3,3m} \\ Y_{31} & Y_{32} & \cdots & Y_{3,3m} \end{bmatrix} \text{ και παίρνουμε:}$$

$$\begin{bmatrix} A_1 & A_2 & A_3 \end{bmatrix} = \begin{bmatrix} \lambda_1^3 Y_{31} & \lambda_2^3 Y_{32} & \cdots & \lambda_{3m}^3 Y_{3,3m} \end{bmatrix} \begin{bmatrix} \lambda_1^2 Y_{31} & \lambda_2^2 Y_{32} & \cdots & \lambda_{3m}^2 Y_{3,3m} \\ \lambda_1 Y_{31} & \lambda_2 Y_{32} & \cdots & \lambda_{3m} Y_{3,3m} \\ Y_{31} & Y_{32} & \cdots & Y_{3,3m} \end{bmatrix}^{-1} \quad (B.19)$$

## B.5 Παρουσίαση των αποτελεσμάτων μετά την τροποποίηση των πόλων

### B.5a Βίντεο 1

Ακολουθούν τα αποτελέσματα για την περίπτωση της φωτιάς:

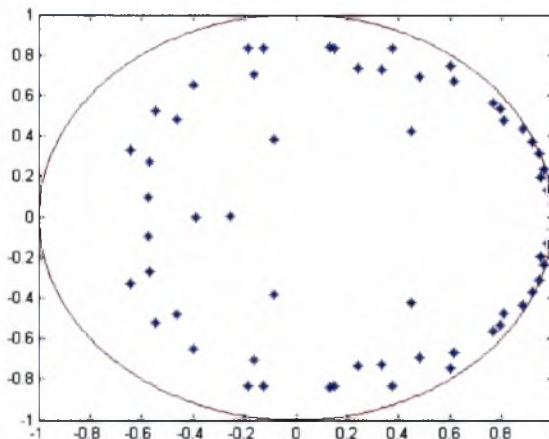




Εικόνα Β.3α. Οι παραπάνω εικόνες αποτελούν το  $150^\circ$  ,  $151^\circ$  ,  $152^\circ$  ,  $153^\circ$  ,  $154^\circ$  ,  $155^\circ$  καρέ αντίστοιχα που συνθέσαμε για μέγεθος διανύσματος κατάστασης ίσο με 20.

Παρατηρώντας το βίντεο διαπιστώνουμε ότι το τελικό αποτέλεσμα δεν διαφέρει σημαντικά από ότι στην περίπτωση που χρησιμοποιούμε μία προηγούμενη κατάσταση. Αυτό ίσως να οφείλεται στο γεγονός ότι για το συγκεκριμένο βίντεο μία προηγούμενη κατάσταση είναι αρκετή για να «συλλάβει» την κίνηση.

Ακολούθως παρουσιάζουμε τους πόλους του συστήματος.



Εικόνα Β.3β. Οι πόλοι του συστήματος.

### Β.5α.1 Χρονική απόδοση του συστήματος κατά το στάδιο της σύνθεσης

Στην παράγραφο αυτή παρουσιάζουμε το χρόνο τον οποίο χρειάζεται το σύστημα για να συνθέσει 200 καρέ. Χρησιμοποιούμε τρία διαφορετικά μήκη διανύσματος κατάστασης.

Μήκος διανύσματος κατάστασης	Χρόνος σύνθεσης
15	4.87
20	5.68
25	6.48

Πίνακας Β.1



Παρατηρούμε ότι οι χρόνοι σύνθεσης είναι ελαφρώς μεγαλύτεροι από ότι στην περίπτωση που χρησιμοποιούμε μία κατάσταση. Αυτό οφείλεται στο γεγονός ότι η μήτρα του συστήματος έχει διάσταση  $n \times 3n$ , σε αντίθεση με την μήτρα του συστήματος που χρησιμοποιεί μία προηγούμενη κατάσταση που έχει διάσταση  $n \times n$ .

### B.5a.2 Χρονική απόδοση του συστήματος κατά το στάδιο της εκμάθησης

Ακολουθεί ο πίνακας στον οποίο παρουσιάζουμε το χρόνο που απαιτείται για να μάθει το μοντέλο την κίνηση. Χρησιμοποιούμε τρία διαφορετικά μήκη διανύσματος κατάστασης (15, 20, 25) και τρία διαφορετικά μεγέθη συνόλου εκμάθησης.

Πλήθος καρέ	Χρόνος εκμάθησης		
	15	20	25
100	8.34	8.36	8.55
90	6.92	6.94	6.97
80	5.47	5.5	5.53

Πίνακας B.2

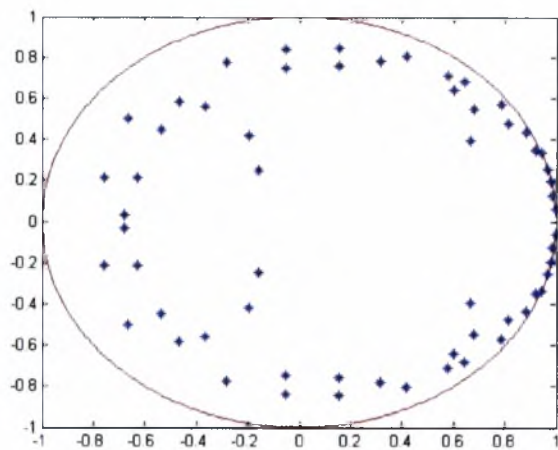
### B.5β Βίντεο 2

Ακολουθούν τα αποτελέσματα που πήραμε για το δεύτερο βίντεο τροποποιώντας τους πόλους που βρίσκονται εκτός μοναδιαίου κύκλου.



Εικόνα B.4a. Οι παραπάνω εικόνες αποτελούν το  $150^\circ$ ,  $151^\circ$ ,  $152^\circ$ ,  $153^\circ$  καρέ αντίστοιχα που συνθέσαμε για μέγεθος διανύσματος κατάστασης ίσο με 20.





Εικόνα Β.4β. Οι πόλοι του συστήματος

## ΜΕΡΟΣ Γ

### Γ.1 Εισαγωγή

Όπως ήδη έχουμε αναφέρει το μοντέλο που χρησιμοποιούμε περιγράφεται μαθηματικά με τις παρακάτω σχέσεις:

$$\begin{aligned}x(n) &= Ax(n-1) + v(n) \\ y(n) &= Cx(n) + w(n)\end{aligned}$$

Σύμφωνα με το μοντέλο αυτό όλα τα καρέ τοποθετούνται σε ένα πίνακα στην RGB μορφή τους και έπειτα γίνεται ο υπολογισμός των διαφορών παραμέτρων.

Στην παρούσα παράγραφο υποθέτω ότι τα καρέ διατίθενται ακολουθιακά και ο υπολογισμός των παραμέτρων θα γίνει σε πραγματικό χρόνο. Συγκεκριμένα για κάθε νέο καρέ που λαμβάνουμε θα υπολογίζουμε όλες τις παραμέτρους του συστήματος και θα συνθέτουμε νέες εικόνες.

### Γ.2 Υπολογισμός της μήτρας του συστήματος

Για τον υπολογισμό της μήτρας  $A$  του συστήματος χρησιμοποιούμε την αναδρομή ελαχίστων τετραγώνων. Αρχικά έχουμε την ποσότητα

$$x(n) - Ax(n-1)$$

την οποία θέλουμε να ελαχιστοποιήσουμε με τη μέθοδο των ελαχίστων τετραγώνων. Δηλαδή:

$$\min_A \sum_{k=1}^n \|X_k - A \cdot X_{k-1}\|^2$$

Παραγωγίζοντας την παραπάνω σχέση ως προς  $A$  παίρνουμε:

$$\begin{aligned}\sum_{k=1}^n X_k \cdot X_{k-1}^T &= \sum_{k=1}^n A \cdot X_k \cdot X_{k-1}^T \Rightarrow \\ \Rightarrow 2 \cdot \sum_{k=1}^n (X_k - A \cdot X_{k-1}) \cdot X_{k-1} &= 0 \Rightarrow \\ \Rightarrow \sum_{k=1}^n X_k \cdot X_{k-1}^T &= A \cdot \sum_{k=1}^n X_{k-1} \cdot X_{k-1}^T \quad (\Gamma.1)\end{aligned}$$

Κατόπιν θέτουμε:

$$R_n = \sum_{k=1}^n X_{k-1} \cdot X_{k-1}^T \quad (\Gamma.2)$$

$$S_n = \sum_{k=1}^n X_k \cdot X_{k-1}^T \quad (\Gamma.3)$$

Ο δείκτης στα  $R_n, S_n$  και  $X_n$  φανερώνει τη χρονική στιγμή στην οποία αναφέρονται οι εν λόγω ποσότητες.

Χρησιμοποιώντας τις δύο τελευταίες σχέσεις η εξίσωση (Γ.1) γίνεται:

$$S_n = A(n) \cdot R_n \quad (\Gamma.4)$$

Πολλαπλασιάζοντας από δεξιά με τον αντίστροφο του πίνακα  $R_n$  παίρνω:

$$A(n) = S_n \cdot R_n^{-1} \quad (\Gamma.5)$$

Η σχέση (Γ.4) για την χρονική στιγμή  $n-1$  γράφεται:

$$S_{n-1} = A(n-1) \cdot R_{n-1} \quad (\Gamma.6)$$

Το  $R_n$  μπορεί να γραφτεί ως:

$$R_n = R_{n-1} + X_{n-1} \cdot X_{n-1}^T \Rightarrow R_{n-1} = R_n - X_{n-1} \cdot X_{n-1}^T \quad (\Gamma.7)$$

Αντίστοιχα το  $S_n$  μπορεί να γραφτεί ως:

$$S_n = S_{n-1} + X_n \cdot X_{n-1}^T \Rightarrow S_{n-1} = S_n - X_n \cdot X_{n-1}^T \quad (\Gamma.8)$$

Αντικαθιστώντας την τελευταία σχέση στην ισότητα (Γ.6) παίρνω:

$$S_n - X_n \cdot X_{n-1}^T = A(n-1) \cdot R_{n-1}$$

Αντικαθιστώντας τώρα την σχέση (Γ.7) στην παραπάνω ισότητα λαμβάνω:

$$S_n - X_n \cdot X_{n-1}^T = A(n-1) \cdot (R_n - X_{n-1} \cdot X_{n-1}^T)$$

Πολλαπλασιάζοντας από δεξιά με  $R_n^{-1}$  παίρνω:

$$\begin{aligned} S_n \cdot R_n^{-1} - X_n \cdot X_{n-1}^T \cdot R_n^{-1} &= A(n-1) \cdot R_n \cdot R_n^{-1} - A(n-1) \cdot X_{n-1} \cdot X_{n-1}^T \cdot R_n^{-1} \Rightarrow \\ S_n \cdot R_n^{-1} &= A(n-1) + X_n \cdot X_{n-1}^T \cdot R_n^{-1} - A(n-1) \cdot X_{n-1} \cdot X_{n-1}^T \cdot R_n^{-1} \end{aligned}$$

Βγάζοντας κοινό παράγοντα το  $X_{n-1}^T \cdot R_n^{-1}$  και χρησιμοποιώντας τη σχέση (Γ.5) η τελευταία ισότητα γίνεται:

$$A(n) = A(n-1) + [X_n - A(n-1) \cdot X_{n-1}] X_{n-1}^T \cdot R_n^{-1} \quad (\Gamma.9)$$

Για να επιτευχθεί η επιθυμητή πολυπλοκότητα  $O(L^2)$  θα πρέπει και ο υπολογισμός του αντιστρόφου  $R_n^{-1}$  να γίνει με αναδρομικό τρόπο.

## Γ.2α Υπολογισμός του αντιστρόφου

Από τη σχέση (Γ.7) έχω:

$$R_n = R_{n-1} + X_{n-1} \cdot X_{n-1}^T \Rightarrow R_n^{-1} = (R_{n-1} + X_{n-1} \cdot X_{n-1}^T)^{-1} \quad (\Gamma.10)$$

Χρησιμοποιώντας την παρακάτω ιδιότητα του αντιστρόφου ενός πίνακα:

$$(A + B \cdot D \cdot C^T)^{-1} = A^{-1} - CA^{-1}(D + C^T A^{-1} B)^{-1} C^T A^{-1}$$

και θέτοντας όπου  $A$  το  $R_{n-1}$ , όπου  $B$  το  $X_{n-1}$ , όπου  $C$  το  $X_{n-1}$  και όπου  $D$  το 1, η σχέση (Γ.10) γίνεται:

$$\begin{aligned} (R_{n-1} + X_{n-1} \cdot 1 \cdot X_{n-1}^T)^{-1} &= R_{n-1}^{-1} - R_{n-1}^{-1} \cdot X_{n-1} \cdot (1 + X_{n-1}^T \cdot R_{n-1}^{-1} \cdot X_{n-1})^{-1} \cdot X_{n-1}^T \cdot R_{n-1}^{-1} = \\ &= R_{n-1}^{-1} - \frac{R_{n-1}^{-1} \cdot X_{n-1} \cdot X_{n-1}^T \cdot R_{n-1}^{-1}}{1 + X_{n-1}^T \cdot R_{n-1}^{-1} \cdot X_{n-1}} \quad (\Gamma.11) \end{aligned}$$

Πολλαπλασιάζοντας την τελευταία σχέση με  $X_{n-1}^T$  από αριστερά παίρνω:

$$\begin{aligned} X_{n-1}^T \cdot R_n^{-1} &= X_{n-1}^T \cdot R_{n-1}^{-1} - \frac{X_{n-1}^T \cdot R_{n-1}^{-1} \cdot X_{n-1} \cdot X_{n-1}^T \cdot R_{n-1}^{-1}}{1 + X_{n-1}^T \cdot R_{n-1}^{-1} \cdot X_{n-1}} = \\ &= \frac{X_{n-1}^T \cdot R_{n-1}^{-1} + X_{n-1}^T \cdot R_{n-1}^{-1} \cdot X_{n-1} \cdot X_{n-1}^T \cdot R_{n-1}^{-1} - X_{n-1}^T \cdot R_{n-1}^{-1} \cdot X_{n-1} \cdot X_{n-1}^T \cdot R_{n-1}^{-1}}{1 + X_{n-1}^T \cdot R_{n-1}^{-1} \cdot X_{n-1}} = \\ &= \frac{1}{1 + X_{n-1}^T \cdot R_{n-1}^{-1} \cdot X_{n-1}} X_{n-1}^T \cdot R_{n-1}^{-1} \quad (\Gamma.12) \end{aligned}$$

Χρησιμοποιώντας τώρα τις δύο σχέσεις:

$$A(n) = A(n-1) + [X_n - A(n-1) \cdot X_{n-1}] X_{n-1}^T \cdot R_{n-1}^{-1} \quad (\Gamma.13)$$

$$X_{n-1}^T \cdot R_n^{-1} = \frac{1}{1 + X_{n-1}^T \cdot R_{n-1}^{-1} \cdot X_{n-1}} X_{n-1}^T \cdot R_{n-1}^{-1} \quad (\Gamma.14)$$

παραθέτουμε τα βήματα του αλγορίθμου και τις αντίστοιχες πολυπλοκότητες:

$$\begin{aligned} K_n &= X_{n-1}^T \cdot R_{n-1}^{-1} & L^2 & \otimes \\ \gamma_n &= \frac{1}{1 + K_n X_{n-1}} & L & \otimes \end{aligned}$$

$$\begin{aligned}
R_n^{-1} &= R_{n-1}^{-1} - \gamma_n K_n^T K_n & 2L^2 & \otimes \\
\varepsilon_n &= X_n - A(n-1)X_{n-1} & L^2 & \otimes \\
A(n) &= A(n-1) + \gamma_n \varepsilon_n K_n & 2L^2 & \otimes
\end{aligned}$$

Συνολικά χρειάζονται  $6L^2$  πολλαπλασιασμούς. Έτσι η πολυπλοκότητα είναι  $O(L^2)$ , ενώ η αρχική πολυπλοκότητα ήταν  $O(L^3)$ .

### Γ.3 Μετασχηματισμός Householder

Θεωρούμε ένα  $m \times 1$  διάνυσμα  $v$  και ένα  $m \times m$  πίνακα ο οποίος ορίζεται ως:

$$H = I - 2 \cdot \frac{v \cdot v^T}{v^T \cdot v} \quad (\Gamma.15)$$

όπου  $I$  ο μοναδιαίος πίνακας διάστασης  $m$ . Το διάνυσμα  $v$  ονομάζεται Householder διάνυσμα και ο πίνακας  $H$  Householder πίνακας.

Ο πίνακας  $H$  έχει κάποιες σημαντικές ιδιότητες:

1. Είναι συμμετρικός

$$H^T = I^T - \frac{2}{v^T \cdot v} (v \cdot v^T)^T = I - \frac{2}{v^T \cdot v} (v^T)^T v^T = I^T - \frac{2}{v^T \cdot v} v \cdot v^T = H \quad (\Gamma.16)$$

2. Είναι ορθογώνιος

$$\begin{aligned}
H^T H &= H H = \left( I - 2 \cdot \frac{v \cdot v^T}{v^T \cdot v} \right) \left( I - 2 \cdot \frac{v \cdot v^T}{v^T \cdot v} \right) = I - 4 \cdot \frac{v \cdot v^T}{v^T \cdot v} + 4 \cdot \frac{v \cdot (v^T \cdot v) \cdot v^T}{(v^T \cdot v)^2} = \\
&= I - 4 \cdot \frac{v \cdot v^T}{v^T \cdot v} + 4 \cdot \frac{v \cdot v^T}{v^T \cdot v} = I \quad (\Gamma.17)
\end{aligned}$$

3. Ισχύει ότι:

$$H H^T = I \Rightarrow H^{-1} = H^T = H \quad (\Gamma.18)$$

4. Αν εφαρμόσουμε τον μετασχηματισμό Householder σε ένα διάνυσμα  $x$ , τότε η ευκλείδεια νόρμα παραμένει σταθερή:

$$\|H \cdot x\|^2 = (H \cdot x)^T H \cdot x = x^T \cdot H^T \cdot H \cdot x$$

αλλά γνωρίζουμε ότι  $H^T \cdot H = I$ . Έτσι η παραπάνω σχέση γίνεται:

$$\|H \cdot x\|^2 = x^T \cdot x = \|x\|^2 \quad (\Gamma.19)$$

### Γ.3α Γεωμετρική ερμηνεία του μετασχηματισμού Householder

Έστω  $x$  ένα διάνυσμα  $m \times 1$  και  $H = I - 2 \cdot \frac{v \cdot v^T}{v^T \cdot v}$  ο πίνακας Householder. Αν εφαρμόσουμε τον μετασχηματισμό στο εν λόγω διάνυσμα θα πάρουμε:

$$H \cdot x = x - 2 \cdot \frac{v \cdot v^T}{v^T \cdot v} \cdot x \quad (\Gamma.20)$$

Γνωρίζουμε όμως ότι  $v^T \cdot x$  είναι βαθμωτό μέγεθος καθώς το διάνυσμα  $v^T$  έχει διαστάσεις  $1 \times m$  και το διάνυσμα  $x$  έχει διαστάσεις  $m \times 1$ . Επίσης γνωρίζουμε ότι  $v^T \cdot v = |v| \cdot |v|$ . Έτσι η παραπάνω σχέση γίνεται:

$$H \cdot x = x - 2 \cdot \frac{v^T \cdot x}{|v|} \cdot \frac{v}{|v|} \quad (\Gamma.21)$$

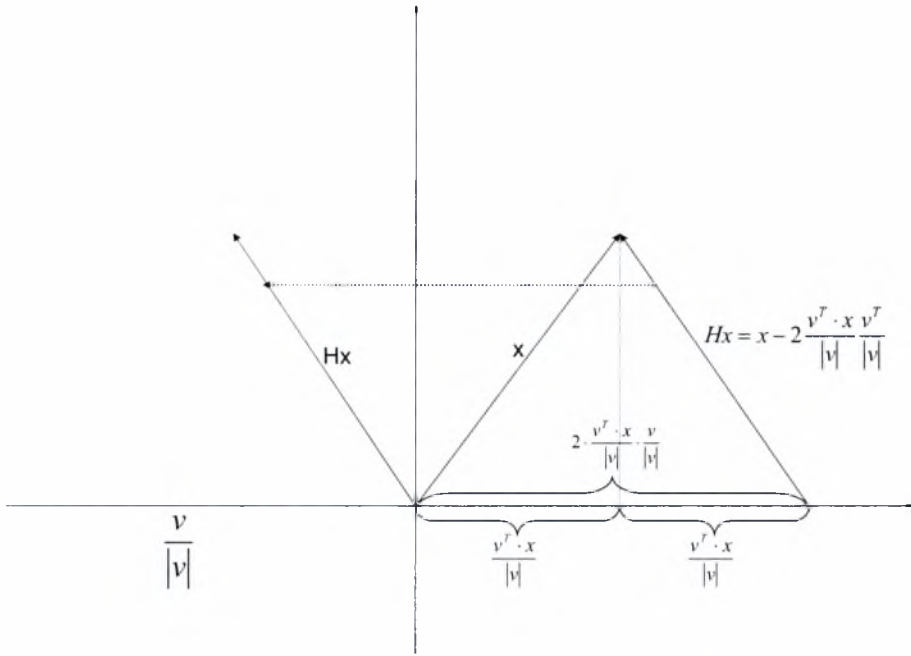
Αλλά το  $\frac{v^T}{|v|}$  όπως και το  $\frac{v}{|v|}$  είναι μοναδιαία διανύσματα καθώς διαιρούμε με το μέτρο τους. Επίσης γνωρίζουμε ότι το εσωτερικό γινόμενο ενός διανύσματος με ένα μοναδιαίο διάνυσμα είναι ίσο με την προβολή του διανύσματος πάνω στο μοναδιαίο διάνυσμα. Στην προκειμένη περίπτωση έχουμε την προβολή του διανύσματος  $x$  πάνω στο μοναδιαίο διάνυσμα  $\frac{v^T}{|v|}$ . Έτσι το διάνυσμα  $2 \cdot \frac{v^T \cdot x}{|v|} \cdot \frac{v}{|v|}$  έχει μέτρο το διπλάσιο της προβολής του  $x$  στη διεύθυνση του  $v$  και διεύθυνση τη διεύθυνση του  $v$ .

Το διάνυσμα τώρα :

$$H \cdot x = x - 2 \cdot \frac{v^T \cdot x}{|v|} \cdot \frac{v}{|v|}$$

θα είναι συμμετρικό του  $x$  ως προς τον άξονα που είναι κάθετος στο  $v$ .

Ακολουθεί η γραφική αναπαράσταση του μετασχηματισμού Householder:



### Γ.3β Εύρεση του διανύσματος Householder

Θα θέλαμε να βρούμε ένα διάνυσμα  $v$  τέτοιο ώστε το  $H \cdot x$  να προβάλλεται

πάνω στον άξονα συντεταγμένων, δηλαδή  $Hx = ae_1$ , όπου  $e_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$  και  $a$  ένα

βαθμωτό μέγεθος. Λύνοντας την σχέση  $H \cdot x = x - 2 \cdot \frac{v \cdot v^T}{v^T \cdot v} \cdot x$  ως προς  $v$  και

θέτοντας  $k = 2 \cdot \frac{v^T \cdot x}{v^T \cdot v}$  παίρνουμε:

$$H \cdot x = x - k \cdot v = ae_1 \Rightarrow v = \frac{x - ae_1}{k} \quad (\Gamma.22)$$

Επιλέγοντας  $k = 1$  (το  $\|v\|$  δεν έχει επίδραση στο μετασχηματισμό) η παραπάνω σχέση γίνεται:

$$v = x - ae_1 \quad (\Gamma.23)$$

Ξέρουμε ότι  $Hx = ae_1 \Rightarrow \|Hx\|_2 = |a|$ , αλλά από την τέταρτη ιδιότητα του μετασχηματισμού ξέρουμε ότι  $\|Hx\|_2 = \|x\|_2$ . Έτσι  $a = \pm \|x\|_2$ . Τελικά το διάνυσμα που παίρνουμε είναι:

$$v = x - \|x\|_2 e_1 \quad (\Gamma.24)$$

## Γ.4 Υπολογισμός της βάσης του συστήματος

Ο υπολογισμός της βάσης του συστήματος θα γίνεται προσαρμοστικά. Αυτό σημαίνει ότι κάθε φορά που θα λαμβάνουμε ένα νέο διάνυσμα, έστω  $y(n) \in \mathfrak{R}^N$ , θα γίνεται νέα εκτίμηση της βάσης. Στην δική μας περίπτωση το διάνυσμα θα αποτελείται από τις RGB τιμές του νέου καρέ. Η μέθοδος που θα ακολουθήσουμε ονομάζεται FDPMP (Fast Data Projection Method).

Η καλύτερη εκτίμηση της βάσης γίνεται κάνοντας ανάλυση ιδιαιζουσών τιμών στη μήτρα συνδιασποράς  $R = E[y(n) \cdot y(n)^T]$  και χρησιμοποιώντας ως βάση τα  $L$  ιδιάζοντα διανύσματα του αριστερού ιδιάζοντος πίνακα που αντιστοιχούν στις  $L$  μεγαλύτερες ιδιάζουσες τιμές. Όμως η υπολογιστική πολυπλοκότητα του svd είναι  $O(N^3)$ , γεγονός που το καθιστά μη εφαρμόσιμο σε εφαρμογές πραγματικού χρόνου. Έτσι αναζητήσαμε άλλους τρόπους για την εκτίμηση της βάσης και αποφασίσαμε να χρησιμοποιήσουμε τη μέθοδο FDPMP, την οποία θα παρουσιάσουμε συνοπτικά παρακάτω.

### Γ.4α Παρουσίαση της μεθόδου FDPMP

Έστω ένας μη αρνητικά ορισμένος πίνακας  $R$ . Από την αριθμητική ανάλυση έχουμε ένα αποτέλεσμα σύμφωνα με το οποίο η βάση που επιθυμούμε να υπολογίσουμε είναι:

$$U_L(n) = \text{orthonormalize}[(I + \mu R)U_L(n-1)] \quad (\Gamma.25)$$

όπου  $\mu$  μία μικρή θετική παράμετρος.

Αν δεν γνωρίζουμε τον πίνακα  $R$ , αλλά διατίθενται ακολουθιακά τα διανύσματα  $y(n) \in \mathfrak{R}^N$ , τότε μπορούμε να τον αντικαταστήσουμε με μία εκτίμησή του που τη χρονική στιγμή  $n$  είναι  $R(n)$  και ισχύει  $E[R(n)] = R$ . Έτσι η παραπάνω σχέση γίνεται:

$$U_L(n) = \text{orthonormalize}[(I + \mu R(n))U_L(n-1)] \quad (\Gamma.26)$$

Χρησιμοποιώντας την στιγμιαία εκτίμηση του  $R(n)$ , δηλαδή  $R(n) = y(n)y(n)^T$  έχουμε:

$$U_L(n) = \text{orthonormalize}[(I + \mu y(n)y(n)^T)U_L(n-1)] \quad (\Gamma.27)$$

Η μέθοδος FDPMP χρησιμοποιεί τον μετασχηματισμό Householder στην διαδικασία της ορθοκανονικοποίησης. Ακολουθεί ο αλγόριθμος FDPMP:



$$\begin{aligned}
 r(n) &= U_L^T(n-1)y(n) \\
 T(n) &= U_L(n-1) + \frac{\bar{\mu}}{\|y(n)\|^2} y(n)r^T(n) \\
 a(n) &= r(n) - \|r(n)\|e_1 \\
 Z(n) &= T(n) - \frac{2}{\|a(n)\|^2} [T(n)a(n)]a(n)^T \\
 U_L(n) &= \text{normalize}(Z(n))
 \end{aligned}$$

### Γ.5 Παρουσίαση αποτελεσμάτων

Στο προηγούμενο εδάφιο που παρουσιάσαμε την μέθοδο FDPM υποθέσαμε ότι μας δίνεται μία αρχική βάση η οποία διαδραματίζει σημαντικό ρόλο στην σύγκλιση του αλγορίθμου. Χρησιμοποιούμε τρεις διαφορετικές αρχικές βάσεις. Παρατηρούμε ότι με τις δύο πρώτες βάσεις ο αλγόριθμος δεν συγκλίνει, κάτι που δεν συμβαίνει με την τρίτη βάση.

Η πρώτη βάση που χρησιμοποιούμε για να αρχικοποιήσουμε τον αλγόριθμο έχει την παρακάτω μορφή:

$$\begin{bmatrix}
 1 & 0 & \dots & 0 \\
 0 & \ddots & & \vdots \\
 \vdots & & \ddots & \vdots \\
 0 & \dots & \dots & 1 \\
 \vdots & \vdots & \vdots & \vdots \\
 1 & 0 & \dots & 0 \\
 0 & \ddots & & \vdots \\
 \vdots & & \ddots & \vdots \\
 0 & \dots & \dots & 1
 \end{bmatrix}$$

Κατόπιν κανονικοποιούμε την βάση ως προς κάθε στήλη.

Για την προσομοίωση χρησιμοποιήσαμε 100 καρέ και μέγεθος διανύσματος κατάστασης ίσο με 35. Παρακάτω παρουσιάζουμε έξι διαδοχικά καρέ. Ο υπολογισμός των παραμέτρων έγινε αφότου λάβαμε και την 100<sup>η</sup> εικόνα.

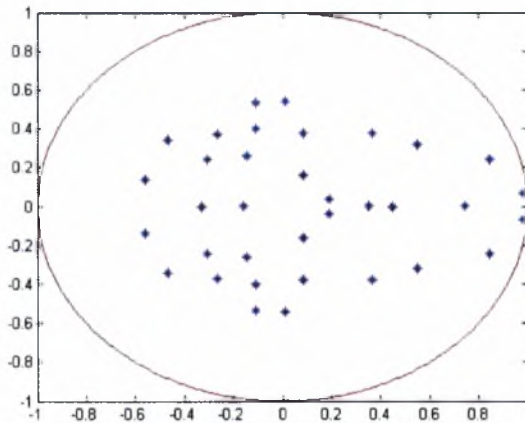




Εικόνα Γ.1α. Οι παραπάνω εικόνες αποτελούν το 20°, 21°, 22°, 23°, 24° και 25° καρέ που συνθέσαμε για μέγεθος διανύσματος κατάστασης ίσο με 35.

Παρατηρούμε ότι η φωτιά δεν παρουσιάζει σημαντικές αλλαγές μεταξύ διαδοχικών καρέ κάτι που οφείλεται στο γεγονός ότι ο αλγόριθμος FDPM δεν συγκλίνει σε 100 καρέ και έτσι θα πρέπει να επιλέξουμε μία διαφορετική βάση για να αρχικοποιήσουμε τον αλγόριθμο.

Ακολουθεί το γράφημα στο οποίο παρουσιάζονται οι πόλοι του συστήματος:



Εικόνα Γ.1β. Πόλοι του συστήματος

Η δεύτερη βάση που επιλέξαμε έχει την παρακάτω μορφή:

$$\begin{bmatrix} 1 & 0 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 1 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}$$

Κατόπιν κανονικοποιούμε την βάση ως προς κάθε στήλη.

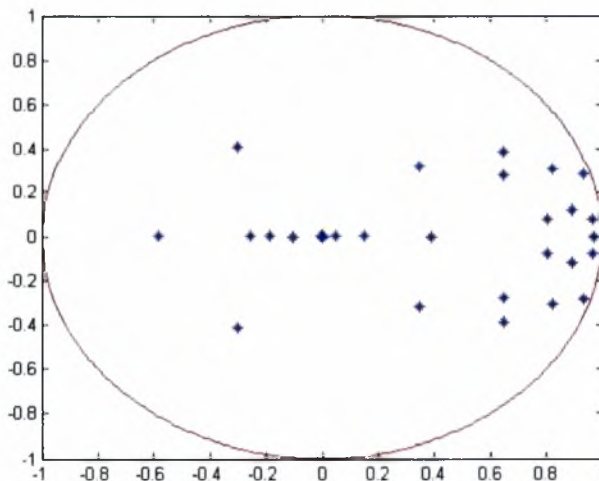
Ακολουθούν τα αποτελέσματα της προσομοίωσης:



Εικόνα Γ.2α. Οι παραπάνω εικόνες αποτελούν το  $20^\circ$ ,  $21^\circ$ ,  $22^\circ$ ,  $23^\circ$ ,  $24^\circ$  και  $25^\circ$  καρέ που συνθέσαμε για μέγεθος διανύσματος κατάστασης ίσο με 35.

Όπως στην πρώτη περίπτωση έτσι και τώρα παρατηρούμε η διαφορά της φωτιάς μεταξύ διαδοχικών εικόνων είναι πολύ μικρή και το πρόβλημα εντοπίζεται πάλι στην εκτίμηση της βάσης.

Ακολουθούν οι πόλοι του συστήματος:



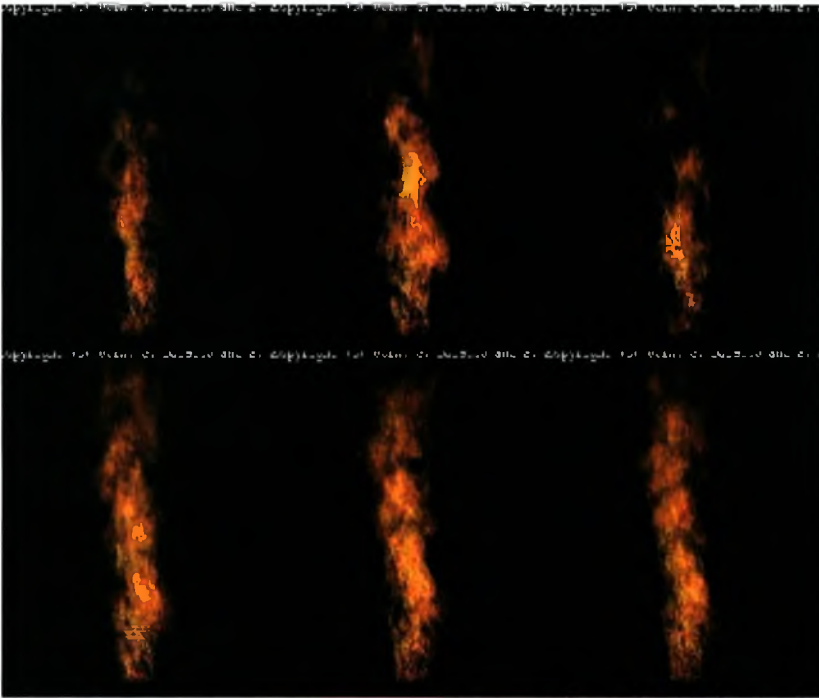
Εικόνα Γ.2β. Πόλοι του συστήματος

Την τρίτη βάση που χρησιμοποιούμε για να αρχικοποιήσουμε τον αλγόριθμο FDPM την υπολογίζουμε με τον τρόπο που παρουσιάσαμε στο πρώτο μέρος της εργασίας, δηλαδή τοποθετούμε τα πρώτα καρέ σε ένα πίνακα στην RGB μορφή τους, αφαιρούμε τον μέσο όρο και παίρνουμε το svd του εν λόγω πίνακα. Οι πρώτες στήλες του αριστερού ιδιάζοντος πίνακα αποτελούν την αρχική μας βάση. Κατόπιν καθώς λαμβάνουμε νέα καρέ εφαρμόζουμε τον αλγόριθμο FDPM, κάνουμε δηλαδή μία νέα εκτίμηση της βάσης.

Στις προσομοιώσεις χρησιμοποιούμε μέγεθος διανύσματος κατάστασης ίσο με 25. Τα καρέ που θα παρουσιάσουμε τα συνθέσαμε αφότου λάβαμε 50 και 100 εικόνες.).

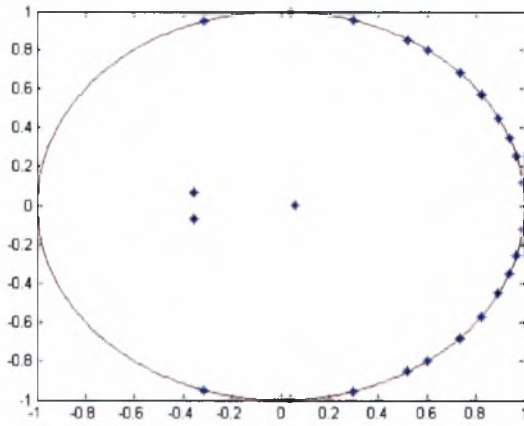
### Γ.5α Βίντεο 1

Ακολουθούν τα αποτελέσματα για μέγεθος διανύσματος κατάστασης ίσο με 25 και αριθμό καρέ ίσο με 50.



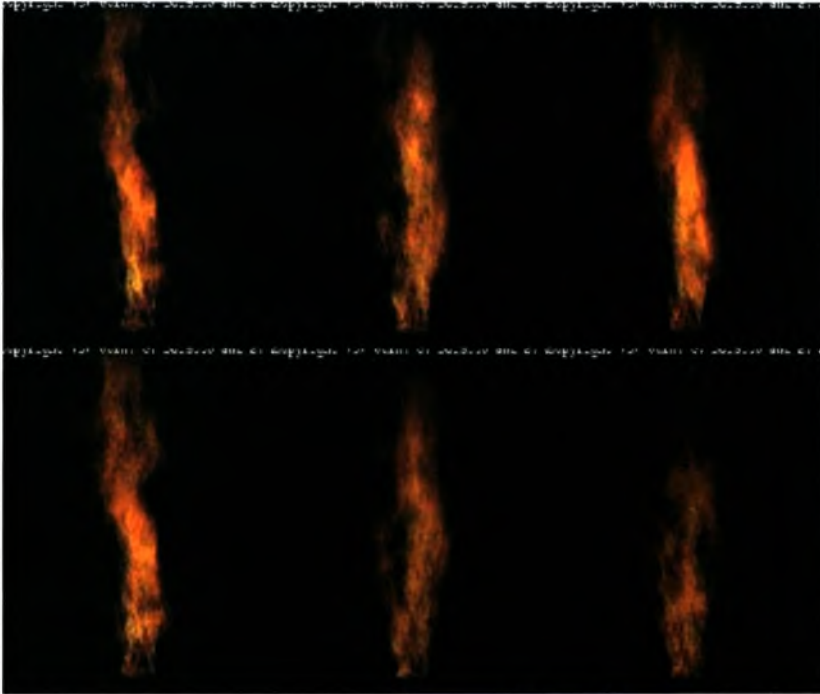
Εικόνα Γ.3α. Οι παραπάνω εικόνες αποτελούν το 20°, 50°, 120°, 200°, 400°, 500° καρέ αντίστοιχα που συνθέσα για μέγεθος διανύσματος κατάστασης ίσο με 25 και αριθμό καρέ ίσο με 50.

Ακολουθεί το διάγραμμα με τους πόλους του συστήματος:

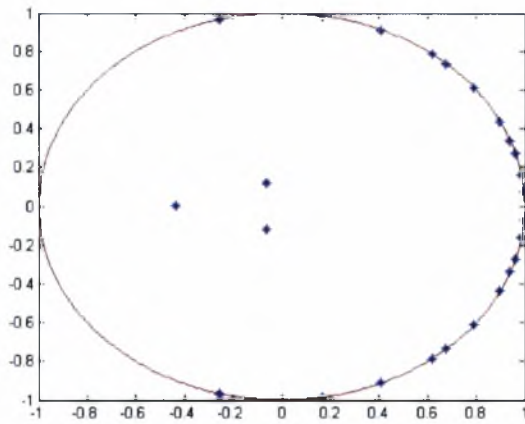


Εικόνα Γ.3β. Οι πόλοι του συστήματος για μέγεθος διανύσματος κατάστασης ίσο με 25 και αριθμό καρτέ ίσο με 50.

Παρακάτω θα παρουσιάσουμε τα αποτελέσματα της προσομοίωσης για μέγεθος διανύσματος κατάστασης ίσο με 25 και αριθμό καρτέ ίσο με 100.



Εικόνα Γ.4α. Οι παραπάνω εικόνες αποτελούν το  $20^\circ$ ,  $50^\circ$ ,  $120^\circ$ ,  $200^\circ$ ,  $400^\circ$ ,  $500^\circ$  καρτέ αντίστοιχα που συνθέσα για μέγεθος διανύσματος κατάστασης ίσο με 25 και αριθμό καρτέ ίσο με 100.



Εικόνα Γ.4β. Οι πόλοι του συστήματος για μέγεθος διανύσματος κατάστασης ίσο με 25 και αριθμό καρτέ ίσο με 100.

Όταν χρησιμοποιούμε 100 καρτέ τα καρτέ που συνθέτουμε είναι ελαφρώς «καλύτερα» από τα αντίστοιχες εικόνες που παίρνουμε όταν χρησιμοποιούμε 50 καρτέ. Αυτό συμβαίνει γιατί στη δεύτερη περίπτωση το σύστημα έχει περισσότερα καρτέ (δηλαδή περισσότερη πληροφορία) για να μάθει την κίνηση της φωτιάς.

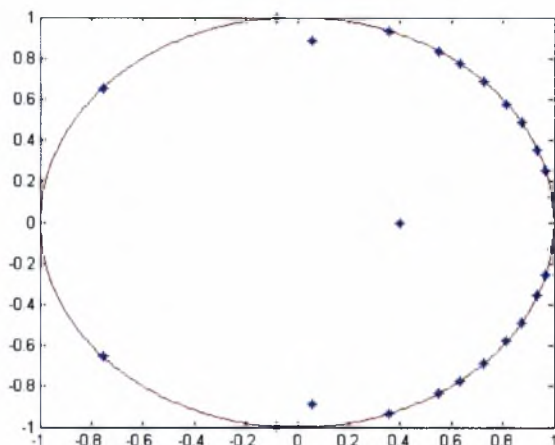
### Γ.5β Βίντεο 2

Ακολουθούν τα αποτελέσματα για μέγεθος διανύσματος κατάστασης ίσο με 25 και αριθμό καρτέ ίσο με 50.





Εικόνα Γ.5α. Οι παραπάνω εικόνες αποτελούν το  $20^\circ$ ,  $100^\circ$ ,  $300^\circ$ ,  $500^\circ$  καρέ αντίστοιχα που συνθέσα για μέγεθος διανύσματος κατάσταση ίσο με 25 και αριθμό καρέ ίσο με 50.

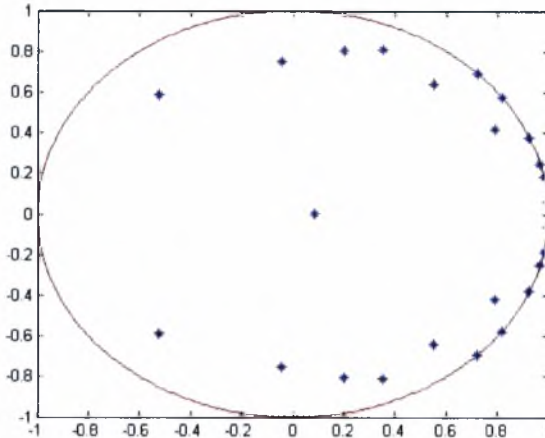


Εικόνα Γ.5β. Οι πόλοι του συστήματος για μέγεθος διανύσματος κατάσταση ίσο με 25 και αριθμό καρέ ίσο με 50.

Οι εικόνες που παίρνουμε για μέγεθος διανύσματος κατάσταση ίσο με 25 και αριθμό καρέ ίσο με 100 είναι:



Εικόνα Γ.6α. Οι παραπάνω εικόνες αποτελούν το 20°, 100°, 300°, 500° καρέ αντίστοιχα που συνθέσα για μέγεθος διανύσματος κατάσταση ίσο με 25 και αριθμό καρέ ίσο με 100.



Εικόνα Γ.6β. Οι πόλοι του συστήματος για μέγεθος διανύσματος κατάσταση ίσο με 25 και αριθμό καρέ ίσο με 100.

Όταν το σύνολο εκμάθησης έχει μέγεθος 100 το βίντεο που συνθέτουμε είναι εμφανώς καλύτερο σε σύγκριση με το βίντεο που συνθέτουμε όταν το σύνολο εκμάθησης έχει μέγεθος 50. Αυτό όπως προανέφερα και στην περίπτωση της φωτιάς οφείλεται στο γεγονός ότι με 50 καρέ περισσότερα το σύστημα είναι σε θέση να συλλάβει κίνηση που χωρίς αυτές τις εικόνες δεν θα ήταν σε θέση να συλλάβει.

## Γ.6 Χρονική απόδοση του συστήματος

Όπως αναφέραμε και στην αρχή του τρίτου μέρους οι συγκεκριμένοι αλγόριθμοι χρησιμοποιούνται σε εφαρμογές πραγματικού χρόνου. Έτσι είναι πολύ σημαντικό να παρουσιάσουμε την χρονική τους απόδοση για διαφορετικά μήκη διανύσματος κατάσταση.

Ακολουθεί ο πίνακας στον οποίο παρουσιάζεται ο χρόνος εκτέλεσης του αλγορίθμου FDPM όταν λαμβάνουμε ένα νέο καρέ.

Μέγεθος διανύσματος κατάσταση	Χρόνος εκτέλεσης (sec)
25	0.344
30	0.406
35	0.469
40	0.532

Πίνακας Γ.1.

Παρατηρούμε ότι ο χρόνος εκτέλεσης του αλγορίθμου αυξάνεται καθώς αυξάνουμε το μήκος του διανύσματος κατάσταση. Πρέπει βέβαια να σημειώσουμε ότι ο χρόνος εκτέλεσης του FDPM είναι σημαντικά μικρότερος από τον χρόνο που απαιτείται για να εφαρμόσουμε svd. Ενδεικτικά αναφέρουμε ότι ο χρόνος εκτέλεσης του svd για πλήθος εικόνων ίσο με 50 είναι 2.33 δευτερόλεπτα.



Ο χρόνος εκτέλεσης της αναδρομής ελαχίστων τετραγώνων και ο χρόνος εκτέλεσης του αλγορίθμου για τον υπολογισμό των υπόλοιπων παραμέτρων του συστήματος χρησιμοποιώντας ακρίβεια τριών δεκαδικών ψηφίων είναι 0.03 δευτερόλεπτα ο κάθε ένας για μέγεθος διανύσματος κατάστασης από 25 έως 40.

Αθροίζοντας τους παραπάνω χρόνους παίρνουμε τον ακόλουθο πίνακα που αφορά το συνολικό χρόνο υπολογισμού των παραμέτρων όταν λαμβάνουμε ένα νέο καρέ.

Μέγεθος διανύσματος κατάστασης	Χρόνος εκτέλεσης (sec)
25	0.404
30	0.466
35	0.529
40	0.592

Πίνακας Γ.2.

Ο τρόπος με τον οποίο συνθέτουμε νέες εικόνες είναι ίδιος με αυτόν που παρουσιάσαμε στο πρώτο μέρος της εργασίας και συνεπώς δεν αλλάζουν οι χρόνοι σύνθεσης.

## Βιβλιογραφία

- [1] Gianfranco Doretto, Alessandro Chiuso, Ying N. Wu, Stefano Soatto, “Dynamic Textures”.
- [2] Gianfranco Doretto, Daniel Cremers, Paolo Favaro, Stefano Soatto, “Dynamic Texture Segmentation”.
- [3] Martin Szummer, Rosalind W. Picard  
“Temporal Texture Modeling”.
- [4] Neill Campbell, Colin Dalton, David Gibson, Barry Thomas  
“Practical Generation of Video Textures using the Auto-Regressive Process”.
- [5] Dmitry Chetverikov, Renaud Peteri  
“A brief survey of dynamic texture description and recognition”.
- [6] Chin-Hwee Peh, Loong-Fah Cheong  
“Synergizing Spatial and Temporal Texture”.
- [7] Lu Yuan, Fang Wen, Ce Liu, Heung-Yeung Shum  
“Synthesizing Dynamic Texture with Closed-loop Linear Dynamic System”.
- [8] Che-Bin Liu, Ruei-sung Lin, Narendra Ahuja  
“Modeling Dynamic Textures Using Subspace Mixtures”.
- [9] Gianfranco Doretto, Stefano Soatto  
“Editable Dynamic Textures”.
- [10] Γεώργιος Β. Μουστακίδης  
“Βασικές Τεχνικές Ψηφιακής Επεξεργασίας Σημάτων”.
- [11] Γεώργιος Β. Μουστακίδης  
“Εισαγωγή στη θεωρία Εκτίμησης και Ανίχνευσης”.
- [12] Xenofon G. Doukopoulos, George V. Moustakides  
“The Fast Data Projection Method For Stable Subspace Tracking”.
- [13] Αθανάσιος Παπούλης  
“Πιθανότητες, Τυχαίες Μεταβλητές Και Στοχαστικές Διαδικασίες”.
- [14] Σέργιος Θεοδωρίδης, Κώστας Μπερμπερίδης  
“Εισαγωγή Στη Θεωρία Σημάτων Και Συστημάτων”.
- [15] Gilbert Strang  
“Γραμμική Άλγεβρα Και Εφαρμογές”.
- [16] Tom M. Mitchell  
“Reducing Data Dimension”.

- [17] Lindsay I Smith  
“A tutorial on Principal Components Analysis”.
- [18] Derek Rowell  
“State-Space Representation of LTI Systems”.
- [19] Julius O. Smith  
“Introduction to State Space Models”.
- [20] <http://www.wikipedia.org/>
- [21] <http://mathworld.wolfram.com/>

ΠΑΝΕΠΙΣΤΗΜΙΟ  
ΘΕΣΣΑΛΙΑΣ



004000085905

