

# Επιταχυντής Υλικού για Συμπίεση Τρισδιάστατων Εικόνων σε Πραγματικό Χρόνο

Δ. Χαϊκάλης, Δ. Μαρούλης, Ν.Σγούρος, Π. Παπαγέωργας

Εθνικό και Καποδιστριακό Πανεπιστήμιο Αθηνών, Τμήμα Πληροφορικής και Τηλεπικοινωνιών  
Πανεπιστημιούπολη, 15784 Ιλίσια  
rtsimage@di.uoa.gr

Περίληψη - Το ενδιαφέρον για τρισδιάστατες εφαρμογές αυξάνει τα τελευταία χρόνια κυρίως χάρη στην αύξηση της διαθέσιμης υπολογιστικής ισχύος και της ανάγκης για ρεαλισμό σε πολλές εφαρμογές, όπως π.χ. στην τηλεόραση, στην ιατρική κ.λπ. Τέτοιες εφαρμογές χειρίζονται εικόνες υψηλής ανάλυσης η οποία οδηγεί σε μεγάλο όγκο δεδομένων για επεξεργασία και μεταφορά. Γι' αυτό το λόγο χρειάζεται ένα ισχυρό σύστημα συμπίεσης, το οποίο θα εκμεταλλεύεται τα ιδιαίτερα χαρακτηριστικά των τρισδιάστατων εικόνων. Η προτεινόμενη υλοποίηση συμπίεσης σε υλικό (hardware) αντιμετωπίζει με επιτυχία τις μεγάλες απαιτήσεις σε εύρος δεδομένων που ένα τέτοιο σύστημα χειρίζεται και κατορθώνει να επιταχύνει τη συγκεκριμένη διαδικασία κατά 1000 φορές σε σχέση με μια αντίστοιχη εφαρμογή σε λογισμικό. Τα αποτελέσματα δείχνουν πως η μεθοδολογία μας μπορεί να εφαρμοστεί σε video εφαρμογές σε συνθήκες πραγματικού χρόνου.

## 1. Εισαγωγή

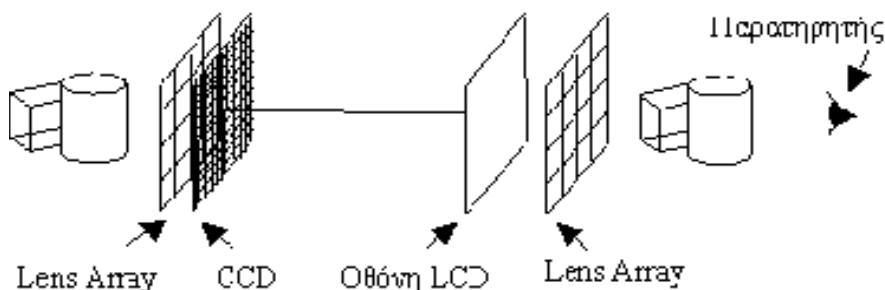
Η διδιάστατη τεχνολογία απεικόνισης χρησιμοποιείται καθημερινά σε μια πληθώρα εφαρμογών. Παρ' όλ' αυτά, υπάρχουν αρκετές εφαρμογές που θα επωφεληθούν από μεγαλύτερο βαθμό ρεαλισμού αν χρησιμοποιηθεί τρισδιάστατη (3D) απεικόνιση. Η αντίληψη του βάθους είναι φυσικό χαρακτηριστικό της όρασης και η τρισδιάστατη απεικόνιση ενισχύει την εντύπωση της παρουσίας του παρατηρητή στη διαδραματιζόμενη σκηνή, σε οποιαδήποτε εφαρμογή [1].

Οι τρισδιάστατες τεχνολογίες απεικόνισης χρειάζονται πρωτόπορους αλγορίθμους και υψηλή υπολογιστική ισχύ για να χειριστούν το μεγάλο όγκο δεδομένων εικόνας. Τα τελευταία χρόνια, η ταχεία αύξηση της υπολογιστικής ισχύος των επεξεργαστών και των επιταχυντών των καρτών γραφικών στους υπολογιστές, συνδυαζόμενα με βελτιώσεις σε οπτικά συστήματα

υψηλής πιστότητας, αναζωογόνησε το ενδιαφέρον για τις 3D εφαρμογές. Πολλές υποσχόμενες τεχνολογίες εξελίχθηκαν, από γυαλιά πόλωσης μέχρι ειδικά γυαλιά με κλείστρα [2] και πιο τελευταία αυτοστερεοσκοπικές συσκευές απεικόνισης [3].

Οι αυτοστερεοσκοπικές συσκευές απεικόνισης προσφέρουν τρισδιάστατη στερεοσκοπική θέαση χωρίς την ανάγκη επιπλέον φορητού εξοπλισμού από το χρήστη. Με αυτό τον τρόπο μειώνεται η κούραση των ματιών, ενώ πολλές από αυτές επιτρέπουν σε περισσότερους από έναν χρήστες να αντιλαμβάνονται την τρισδιάστατη εντύπωση [4]. Στόχος αυτών των συσκευών είναι η ρεαλιστική απεικόνιση της σκηνής μπροστά στο θεατή, αποφεύγοντας φαινόμενα όπως αυτό του «κουκλοθέατρου», κατά το οποίο η αναπαραγωγή προσώπων είναι αφύσικα μικρή [5]. Τέτοιες συσκευές απεικόνισης συνήθως αποτελούνται από μια οθόνη υγρών κρυστάλλων (LCD) υψηλής ανάλυσης εξοπλισμένες με ένα κατάλληλο σύστημα παράλλαξης. Πιο πρόσφατα άρχισαν να κατασκευάζονται οθόνες βασισμένες σε μικροφακούς. Τέτοια συστήματα απεικόνισης είναι χρήσιμα σε ιατρικές [6], εκπαιδευτικές και ψυχαγωγικές [7] εφαρμογές. Μια ειδική κατηγορία αυτοστερεοσκοπικών οθονών που προσφέρει ενισχυμένη αντίληψη του βάθους και υποστηρίζει απεικονίσεις όλων των χρωμάτων, λειτουργεί με τις αρχές της ολοκληρωτικής φωτογραφίας (Integral photography, IP), η οποία πρωτοχρησιμοποιήθηκε από τον Lippman [8] το 1908.

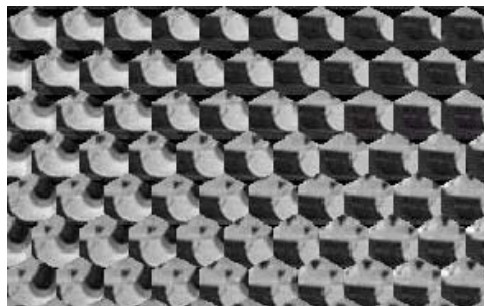
Στην Εικόνα 1 σκιαγραφούνται οι αρχές λήψης και απεικόνισης της ολοκληρωτικής φωτογραφίας. Μεταξύ της διάταξης λήψης και του αντικειμένου παρεμβάλλεται μια συστοιχία μικροφακών (lens array). Κατά την απεικόνιση, μια πανομοιότυπη συστοιχία τοποθετείται μεταξύ του παρατηρητή και της οθόνης απεικόνισης, δημιουργώντας έτσι την αντίληψη του βάθους στην προβολή του αντικειμένου.



Εικόνα 1: Μια τυπική διάταξη λήψης και απεικόνισης ολοκληρωτικής φωτογραφίας.

Οι τρισδιάστατες εφαρμογές απαιτούν τη χρήση υψηλής ανάλυσης εικόνων, η οποία οδηγεί σε μεγάλο εύρος δεδομένων και αντίστοιχες ανάγκες αποθήκευσης για τη λήψη και την αναπαραγωγή 3D αντικειμένων και σκηνών. Επομένως, είναι πρωτίστης σημασίας η χρήση ενός αποτελεσματικού αλγόριθμου συμπίεσης των σχετικών δεδομένων.

Στην Εικόνα 2 φαίνεται ένα μέρος μιας IP εικόνας. Μια τέτοια εικόνα σχηματίζεται από υπο-εικόνες οι οποίες είναι συνήθως διατεταγμένες σε τετράγωνη ή εξαγωνική γεωμετρία, ανάλογα με τη συστοιχία μικροφακών που χρησιμοποιείται σε κάθε περίπτωση. Όπως είναι φανερό, γειτνιάζουσες υπο-εικόνες παρουσιάζουν μεγάλη συσχέτιση η οποία οδηγεί σε μεγάλο όγκο πλεονάζουσας πληροφορίας.



Εικόνα 2: μέρος μιας IP εικόνας. Οι γειτνιάζουσες υπο-εικόνες έχουν έντονο πλεονασμό.

Αυτές οι επίκτητες ιδιότητες μιας IP εικόνας έδωσαν το έναυσμα για δημιουργία νέων τεχνικών για τη συμπίεση των υψηλά συσχετισμένων δεδομένων με σκοπό να υπερβούν την ικανότητα συμπίεσης παραδοσιακών τεχνικών. Οι τεχνικές που χρησιμοποιούνται βασίζονται κυρίως σε υψηλότερων διαστάσεων μετασχηματισμούς συνημιτόνου (DCT), με σημαντικά αποτελέσματα για ένα συγκεκριμένο τύπο IP εικόνων.

Τα προτερήματα μια τέτοιας τεχνικής που χρησιμοποιεί έναν τρισδιάστατο DCT επισκιάζονται από την πολυπλοκότητα και την έλλειψη ισχύος λόγω των πολύπλοκων πινάκων κβάντισης σε πολυδιάστατους χώρους. Μια εναλλακτική τεχνική η οποία πετυχαίνει υψηλούς βαθμούς συμπίεσης για οποιαδήποτε τρισδιάστατη εφαρμογή παρουσιάζεται στο [9] και είναι προϊόν ερευνητικής προσπάθειας μελών της ομάδας μας. Η μέθοδος αυτή χαρακτηρίζεται από μεγάλο βαθμό ρωμαλεότητας και ικανότητας βασιζόμενη στη γενική ιδέα του αλγόριθμου κωδικοποίησης MPEG. Στόχος της μεθόδου είναι οι εφαρμογές πραγματικού χρόνου, όπου υψηλοί λόγοι συμπίεσης πρέπει να συνδυάζονται με εκτέλεση σε πραγματικό χρόνο ολόκληρης της διαδικασίας.

Οι υλοποιήσεις σε Επαναπρογραμματιζόμενες Συστοιχίες Πυλών (FPGAs) εκμεταλλεύονται παράλληλες αρχιτεκτονικές που επιταχύνουν πολύπλοκες εργασίες για τις οποίες ο χρόνος εκτέλεσης διαδραματίζει σημαντικό ρόλο. Τα βασικά προτερήματα των FPGAs συνοψίζονται στη μεταβλητότητα του σχεδιασμού, τους γρήγορους χρόνους σχεδιασμού χάρη στα ειδικά εργαλεία και τις γλώσσες περιγραφής υλικού, και στην ολοένα αυξανόμενη απόδοσή τους [10].

Στις επόμενες παραγράφους, περιγράφεται ο αλγόριθμος συμπίεσης IP εικόνων και εξηγείται ο σχεδιασμός του σε υλικό. Παρατίθενται επίσης και γίνονται συγκρίσεις μεταξύ χρόνων εκτέλεσης του αλγορίθμου συμπίεσης σε λογισμικό και από την κάρτα (FPGA) της Celoxica.

## 2. Περιγραφή αλγορίθμου

### 2.1 Αρχές προτύπου MPEG

Μια από τις πιο κρίσιμες διεργασίες ενός σχήματος συμπίεσης που βασίζεται σε τεχνικές όμοιες με το MPEG είναι η εκτίμηση κίνησης. Μια περιοχή αναζήτησης στο πλαίσιο αναφοράς διασχίζεται με σκοπό να βρεθεί το καλύτερο ταίριασμα μιας ομάδας εικονοστοιχείων (μπλοκ) από το τρέχον πλαίσιο. Στη συνέχεια, οι διαφορές στις τιμές φωτεινότητας των εικονοστοιχείων των μπλοκ και η διαφορά στις συντεταγμένες (διάνυσμα κίνησης) κωδικοποιούνται. Δύο βασικά ζητήματα σχετίζονται με την εκτίμηση κίνησης: το μέγεθος της περιοχής αναζήτησης και το μέτρο που καθορίζει το «καλύτερο ταίριασμα». Και για τα δύο ζητήματα έχουν προταθεί αρκετές μέθοδοι [11] ώστε να μειωθεί ο αριθμός των υπολογισμών.

Το ευρύτερα χρησιμοποιούμενο μέτρο είναι το Άθροισμα των Απολύτων Διαφορών (Sum of Absolute Differences – SAD), το οποίο προσθέτει την απόλυτη τιμή των διαφορών μεταξύ των αντίστοιχων εικονοστοιχείων των συγκρινόμενων μπλοκ. Η διαδικασία του υπολογισμού των τιμών SAD είναι χρονοβόρα λόγω της σύνθετης φύσης του υπολογισμού της απόλυτης τιμής και το πλήθος των προσθέσεων. Για ένα μπλοκ 8x8, η τιμή SAD υπολογίζεται με τον εξής τύπο:

$$SAD(x, y, r, s) = \sum_{i=0}^7 \sum_{j=0}^7 \left( A_{(x+i, y+j)} - B_{((x+r)+i, (y+s)+j)} \right)$$

όπου:  $0 \leq x, y <$  μέγεθος πλαισίου,  $(r, s)$  το διάνυσμα κίνησης,  $A_{(x,y)}$  το εικονοστοιχείο του τρέχοντος πλαισίου στη θέση  $(x, y)$  και  $B_{(x,y)}$  το εικονοστοιχείο του πλαισίου αναφοράς στη θέση  $(x, y)$ .

Ένα δεύτερο ζήτημα είναι ο τύπος του μετασχηματισμού που θα χρησιμοποιηθεί για την κωδικοποίηση των πλαισίων. Ο DCT χρησιμοποιείται σε πολλά συστήματα συμπίεσης πολυμέσων (JPEG, MPEG, H.261) χάρη στη σχετικά χαμηλή πολυπλοκότητά του. Παρά τη σχεδιαστική απλότητά του, ακόμα και ένας τυπικός  $8 \times 8$  διδιάστατος (2D) DCT έχει σημαντικό υπολογιστικό κόστος, αφού χρειάζεται να εκτελέσει 4096 υπολογισμούς ανά παράθυρο [12].

Στις εφαρμογές συμπίεσης πολυμέσων, είναι ευρέως παραδεκτό ότι, εφαρμόζοντας τον DCT σε παράθυρα μεγέθους  $8 \times 8$ , πετυχαίνεται μια καλή ανταλλαγή μεταξύ της πολυπλοκότητας του μετασχηματισμού και της χωρικής συσχέτισης [13]. Η έκφραση για έναν  $8 \times 8$  διδιάστατο DCT (2D-DCT) είναι:

$$F_{m,n} = \sum_{i=0}^7 \sum_{j=0}^7 f_{i,j} \cos\left(\frac{\pi m_i}{16}\right) \cos\left(\frac{\pi n_j}{16}\right)$$

όπου  $f_{i,j}$  είναι τα δείγματα εισόδου,  $F_{m,n}$  είναι το μετασχηματισμένο σήμα και  $m_i = (2i + 1)m$  και  $n_j = (2j + 1)n$  οι συντελεστές.

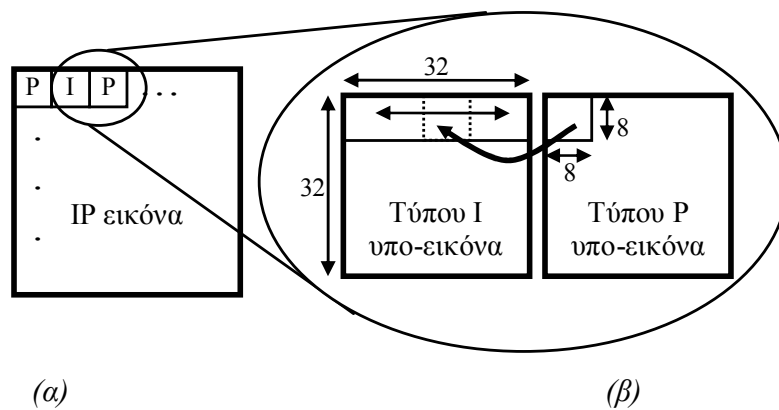
## 2.2 Περιγραφή αλγορίθμου της μεθόδου

Ένα απλό σχήμα κωδικοποίησης MPEG αποτελείται από τυπικούς χάρτες κβάντισης, διαδικασίες κωδικοποίησης, πρόβλεψη διανυσμάτων κίνησης και κωδικοποίηση. Στην εργασία μας [9] η τεχνική συμπίεσης στηρίζεται στο πρότυπο MPEG. Κάποιες μετατροπές όμως γίνονται στους αλγορίθμους εκτίμησης κίνησης ώστε να χρησιμοποιούνται χωρίς υπερβολικό υπολογιστικό κόστος, ενώ παράλληλα να αυξάνεται η ποιότητα της τελικής IP εικόνας.

Στην Εικ. 3 φαίνεται η κατάτμηση μιας IP εικόνας σε υπο-εικόνες και η περιοχή αναζήτησης των P μπλοκ στην I υπο-εικόνα. Η μέθοδος συμπίεσης εφαρμόζεται σε μία IP εικόνα, χρησιμοποιώντας κάθε υπο-εικόνα ως ένα χωρικό πλαίσιο το οποίο είναι υψηλά

συσχετισμένο με τις γειτνιάζουσες υπο-εικόνες. Με αυτό τον τρόπο, τα διανύσματα κίνησης του αλγορίθμου MPEG αντικαθιστούνται από διανύσματα διαφοράς, τα οποία αντιπροσωπεύουν την αλλαγή της σχετικής θέσης των συντεταγμένων του μπλοκ καλύτερου ταιριάσματος μεταξύ δύο υπο-εικόνων που δρουν ως πλαίσιο αναφοράς και τρέχον πλαίσιο.

Ο τύπος αναζήτησης που χρησιμοποιείται είναι εξαντλητική σε μία διεύθυνση, γιατί, όπως αναφέρεται στο [9], ο συγκεκριμένος τύπος οδηγεί στην εύρεση του καλύτερου ταιριάσματος με απόδοση πολύ κοντινή στις περισσότερες τυπικές μεθόδους αναζήτησης. Μια ενδιάμεση (τύπου I) υπο-εικόνα μεγέθους 32x32 συμπιέζεται με τρόπο παρόμοιο με το JPEG, ακολουθούμενος από κβάντιση και κατάλληλη κωδικοποίηση με στόχο την αύξηση της απόδοσης του αλγορίθμου. Οι δύο υπο-εικόνες πρόβλεψης (τύπου P) μεγέθους 32x32 που βρίσκονται εκατέρωθεν της I υπο-εικόνας κατασκευάζονται με τον υπολογισμό των κατάλληλων διανυσμάτων διαφοράς με βάση την ανακατασκευασμένη I υπο-εικόνα. Για τον υπολογισμό αυτό, χρησιμοποιείται το μέτρο SAD το οποίο περιγράφηκε στην παράγραφο 2.1. Η αναζήτηση του καλύτερου ταιριάσματος και ο υπολογισμών των τιμών SAD υλοποιούνται σε FPGA, με στόχο τη λειτουργία σε πραγματικό χρόνο.



Εικόνα 3: (α) Η διάταξη των I και P υπο-εικόνων σε μια IP εικόνα, (β) η περιοχή αναζήτησης των P μπλοκ στην I υπο-εικόνα.

### 3. Σχεδιασμός υλικού

Για την υλοποίηση του αλγορίθμου σε υλικό (hardware), χρησιμοποιείται γλώσσα περιγραφής υλικού VHDL. Το σχέδιο προσομοιώνεται και ελέγχεται με τον προσομοιωτή ModelSim και το τελικό αρχείο προγραμματισμού δημιουργείται με το αναπτυξιακό εργαλείο

Xilinx ISE 5.1. Το τελικό ψηφιακό κύκλωμα ενσωματώνεται σε Xilinx Virtex XVC-2000E FPGA που βρίσκεται στην αναπτυξιακή κάρτα PCI RC1000-PP της Celoxica.

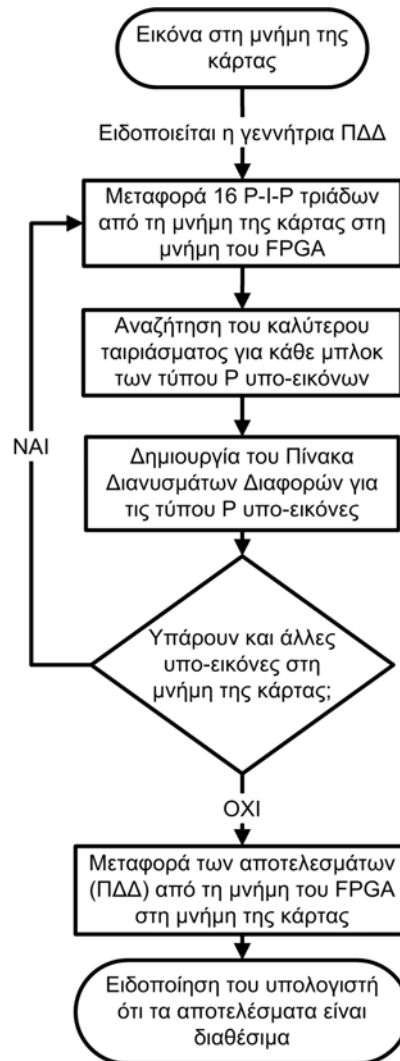
Το συγκεκριμένο FPGA παρέχει ισοδύναμη περιοχή 2 εκατομμυρίων λογικών πυλών. Η κάρτα περιλαμβάνει 8 MB μνήμης και το FPGA ενσωματώνει 80 kbytes μνήμης διπλής προσπέλασης, η οποία μπορεί να χρησιμοποιηθεί για οποιοδήποτε εύρος λέξης [14]. Στην εφαρμογή μας, ένα μέρος μιας IP εικόνας μεγέθους ως 2 MB στέλνεται στη μνήμη της κάρτας και από εκεί μπορούν να διαβαστούν από το ψηφιακό κύκλωμα παράλληλα τρεις υπο-εικόνες που σχηματίζουν P-I-P τριάδες υπο-εικόνων. Το ψηφιακό κύκλωμα δημιουργεί τα διανύσματα διαφοράς, και το ονομάζουμε γεννήτρια Πινάκων Διανυσμάτων Διαφοράς (ΠΔΔ).

Στην Εικ. 4 παρουσιάζεται ένα απλοποιημένο διάγραμμα του κύκλου λειτουργίας της γεννήτριας ΠΔΔ. Η πρώτη λειτουργία της είναι να μεταφέρει ένα μέρος της εικόνας στην εσωτερική μνήμη του. Ο τρόπος με τον οποίο πρέπει να εισαχθούν τα δεδομένα στις μονάδες επεξεργασίας καθορίζει το μέγεθος της εσωτερικής μνήμης που υλοποιείται στη γεννήτρια ΠΔΔ για αποθήκευση. Έτσι, σε κάθε κύκλο λειτουργίας της γεννήτριας ΠΔΔ, 16 P-I-P τριάδες μεταφέρονται στην εσωτερική μνήμη. Η τελευταία είναι διαμορφωμένη σε 4 τράπεζες, τρεις για τις διαφορετικές υπο-εικόνες που αποτελούν μία τριάδα και άλλη μία για την εγγραφή των αποτελεσμάτων.

Από εκεί, μια σειρά υπολογισμών εκτελούνται με σκοπό να δημιουργηθούν και να αποθηκευτούν στη μνήμη των αποτελεσμάτων οι πίνακες που περιέχουν τα διανύσματα διαφοράς για κάθε P υπο-εικόνα. Για τη δημιουργία των διανυσμάτων διαφοράς, χρησιμοποιείται το μέτρο SAD, όπως αναφέρθηκε στην προηγούμενη παράγραφο. Ο υπολογισμός κάθε τιμής SAD περιλαμβάνει 64 συγκρίσεις ζευγών εικονοστοιχείων, αντιστροφές και μια σειρά προσθέσεων. Για να εκτελεστούν οι υπολογισμοί των τιμών SAD εντελώς παράλληλα, πρέπει να χρησιμοποιηθούν περισσότερα από ένα FGPA's [15], λόγω των μεγάλων απαιτήσεων μιας τέτοιας υλοποίησης σε ροή εισόδου/εξόδου δεδομένων. Στην προτεινόμενη μέθοδο, δημιουργούμε ένα σύστημα που μπορεί να λειτουργήσει ενσωματωμένο μόνο σε ένα FPGA, υπολογίζοντας μια τιμή SAD σε 8 κύκλους ρολογιού του συστήματος, εκτελώντας 8 συγκρίσεις σε κάθε κύκλο.

Αφού χρησιμοποιηθούν όλα τα δεδομένα που υπάρχουν στην εσωτερική μνήμη, οι επόμενες 16 P-I-P τριάδες μεταφέρονται εκεί από τη μνήμη της κάρτας, και με αυτό τον τρόπο συντελείται η δημιουργία των πινάκων διανυσμάτων διαφοράς μέχρι να επεξεργαστούν όλα τα

δεδομένα που βρίσκονται στη μνήμη της κάρτας. Στο σημείο αυτό, τα αποτελέσματα που έχουν αποθηκευτεί στην εσωτερική μνήμη αποτελεσμάτων της γεννήτριας ΠΔΔ μεταφέρονται στη μνήμη της κάρτας και ο υπολογιστής ειδοποιείται για να τα παραλάβει.

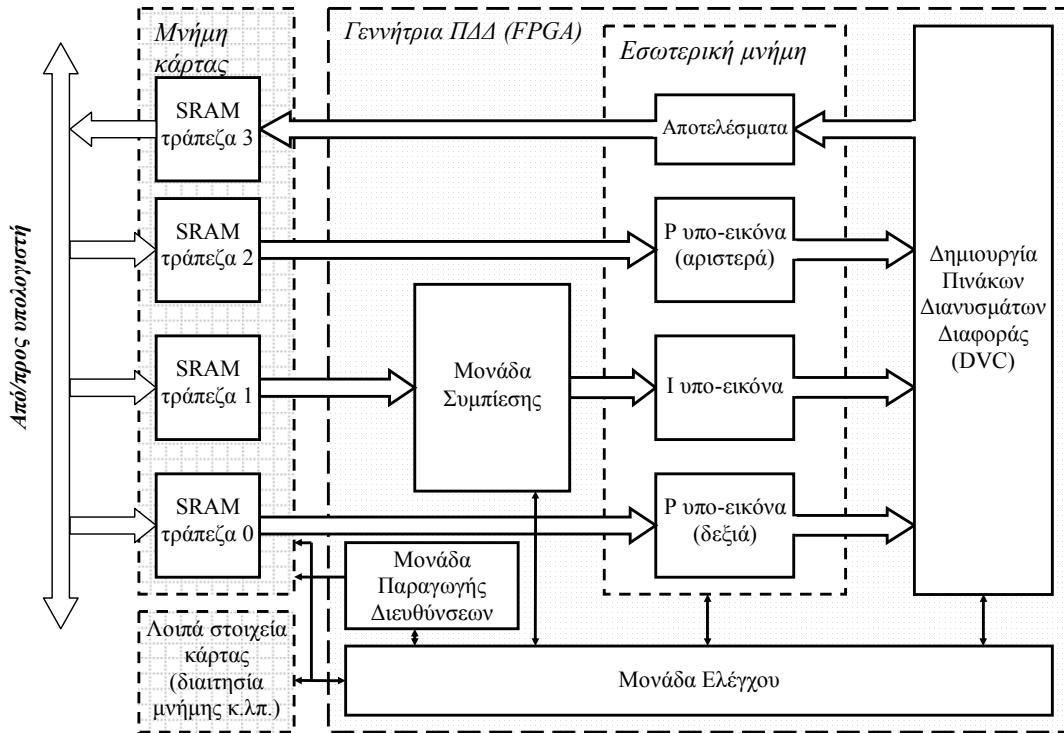


Εικόνα. 4: Διάγραμμα του κύκλου λειτουργίας της γεννήτριας ΠΔΔ.

Η γεννήτρια ΠΔΔ αποτελείται από της εξής υπο-μονάδες: διδιάστατος DCT, κβαντιστής, αντίστροφος κβαντιστής, αντίστροφος 2D-DCT, μονάδα δημιουργίας πινάκων διανυσμάτων διαφοράς (DVC), μονάδα παραγωγής διευθύνσεων (AGU) και μονάδα ελέγχου (CU), καθώς και από τις απαραίτητες μνήμες και τον έλεγχό τους. Οι τέσσερις πρώτες υπο-μονάδες είναι ομαδοποιημένες στη μονάδα Συμπίεσης η οποία είναι υπεύθυνη για τη συμπίεση και



ανακατασκευή των υπο-εικόνων τύπου I. Στην Εικ. 5 φαίνεται ένα σχηματικό διάγραμμα της γεννήτριας ΠΔΔ και των επιμέρους μονάδων που το αποτελούν.



Εικόνα 5: Σχηματική αναπαράσταση του υλοποιημένου ψηφιακού κυκλώματος (γεννήτρια ΠΔΔ)

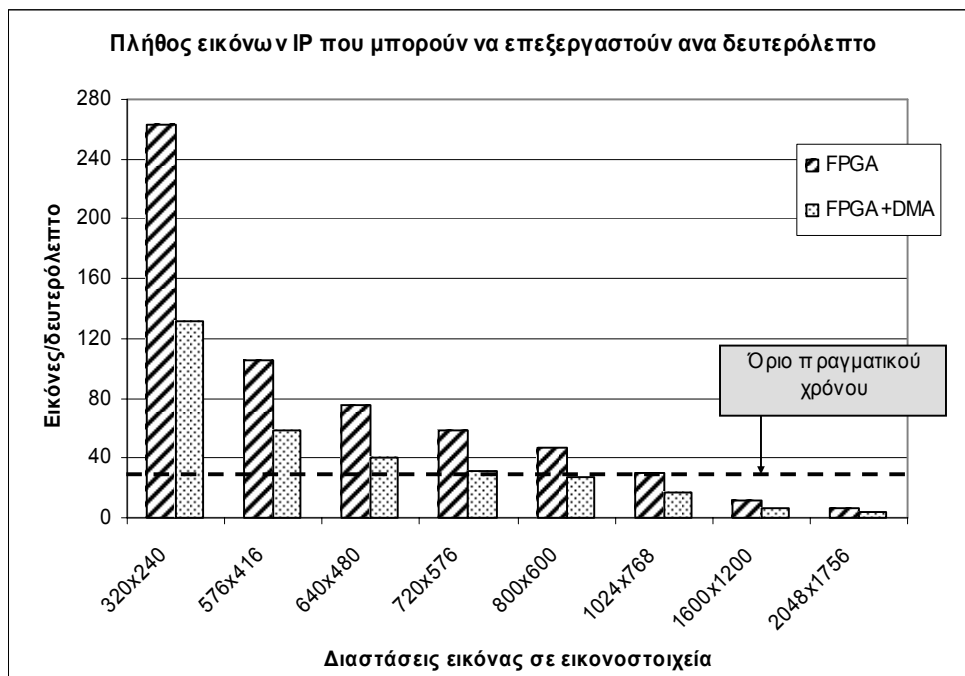
#### 4. Αποτελέσματα

Όπως αναφέρθηκε στην προηγούμενη παράγραφο, η γεννήτρια ΠΔΔ επεξεργάζεται 16 P-I-P τριάδες σε κάθε κύκλο λειτουργίας του. Ο συνολικός χρόνος για αυτό το σύνολο των υπολογισμών προκύπτει από το άθροισμα των επιμέρους χρόνων των λειτουργιών των διαφόρων μονάδων που το αποτελούν. Έτσι, για τη μεταφορά των 16 τριάδων στη μνήμη της γεννήτριας ΠΔΔ, τον υπολογισμό των τιμών SAD, την εύρεση της μικρότερης τιμής, την εγγραφή των αποτελεσμάτων στην αντίστοιχη εσωτερική μνήμη και τη μεταφορά τους από αυτή στη μνήμη της κάρτας, χρειάζονται συνολικά 21.250 κύκλοι ρολογιού.

Το ψηφιακό σύστημα λειτουργεί με μέγιστη συχνότητα λειτουργίας 11,29 MHz και καταλαμβάνει 87% της επιφάνειας του Virtex XVC-2000E FPGA. Η χαμηλή συχνότητα λειτουργίας δικαιολογείται από την έλλειψη εκτεταμένων σταδίων παραλληλισμού σε ορισμένες

μονάδες της γεννήτριας ΠΔΔ λόγω περιορισμών στη διαθέσιμη επιφάνεια υλοποίησης. Με αυτή τη συχνότητα, προκύπτει ότι 1,9 ms χρειάζονται για τη δημιουργία των αποτελεσμάτων για 16 τριάδες υπο-εικόνων. Ο χρόνος επεξεργασίας μιας ολόκληρης IP εικόνας εξαρτάται από το μέγεθός της.

Εκτός από το χρόνο επεξεργασίας του FPGA, πρέπει να ληφθεί υπόψη και ο χρόνος μεταφοράς των δεδομένων από τον υπολογιστή στη μνήμη της κάρτας μέσω καναλιού DMA, αν θεωρήσουμε το ψηφιακό κύκλωμα ως μέρος ενός ολοκληρωμένου συστήματος συμπίεσης IP εικόνων. Στο Διάγραμμα 6, παρουσιάζονται οι χρόνοι επεξεργασίας IP εικόνων διαφόρων διαστάσεων ανά δευτερόλεπτο.



Διαγραμμα 6: ο αριθμός των IP εικόνων τις οποίες μπορεί να επεξεργαστεί η γεννήτρια ΠΔΔ ανά δευτερόλεπτο

Όπως φαίνεται από το διάγραμμα, παρά τη χαμηλή συχνότητα λειτουργίας, η γεννήτρια ΠΔΔ επεξεργάζεται εικόνες σημαντικού μεγέθους, επομένως κρίνεται κατάλληλη η εφαρμογή της σε επεξεργασία κινούμενης εικόνας (video) σε πραγματικό χρόνο.

Λόγω της πρωτοτυπίας της μεθόδου, δε μπορεί να γίνει άμεση σύγκριση χρόνων λειτουργίας και επεξεργασίας με άλλες υλοποιήσεις. Αναφέρουμε ότι πρόσφατη υλοποίηση σε FPGA ενός σχήματος εκτίμησης κίνησης με βελτιστοποίηση του κυκλώματος ως προς την

επιφάνεια και με ελαχιστοποίηση του εύρους της μνήμης [16] πετυχαίνει μέγιστη συχνότητα λειτουργίας 8 MHz, και μπορεί να επεξεργαστεί σε πραγματικό χρόνο εικόνες με μέγεθος μικρότερο από 640x480 εικονοστοιχεία.

Συγκρίνοντας τα αποτελέσματα με μια ανάλογη υλοποίηση σε λογισμικό, η οποία δημιουργήθηκε σε γλώσσα προγραμματισμού C, προκύπτει ότι η γεννήτρια ΠΔΔ εκτελεί τους υπολογισμούς 1000 περίπου φορές γρηγορότερα. Συγκεκριμένα, το λογισμικό χρειάζεται αρκετά δευτερόλεπτα για να δημιουργήσει τους πίνακες με τα διανύσματα διαφοράς, ακόμα και για εικόνες μικρών διαστάσεων. Ο κώδικας του προγράμματος δημιουργήθηκε και μεταγλωττίστηκε με το πρόγραμμα Microsoft Visual Studio 6 και οι χρόνοι του μετρήθηκαν σε υπολογιστή Pentium 4 στα 2400 MHz, με μνήμη συστήματος 512MB. Ο Πίνακας 7 παρουσιάζει τη σύγκριση των παραπάνω χρόνων φανερώνοντας το πλεονέκτημα της γεννήτριας ΠΔΔ:

		Τρόπος υπολογισμού		
		FPGA (ms)	FPGA + μεταφορά DMA (ms)	Λογισμικό (ms)
Διαστάσεις εικόνας σε εικονοστοιχεία	576x416	9.5	16.9	18.300
	800x600	20.9	36.6	40.100
	1024x768	32.3	56.4	61.800

Πίνακας 7: χρόνοι εκτέλεσης σε υλικό και λογισμικό του υπολογισμού ΠΔΔ μιας IP εικόνας.

## 5. Συμπεράσματα

Το ολοένα και αυξανόμενο ενδιαφέρον για τρισδιάστατες εφαρμογές έχει οδηγήσει στην αναβίωση της έρευνας για μεθόδους και τεχνικές λήψης και αναπαραγωγής 3D εικόνων. Τα σημερινά μέσα επεξεργασίας και απεικόνισης πρέπει να χειριστούν μεγάλο όγκο δεδομένων που προκύπτει από εικόνες υψηλής ανάλυσης. Στην εργασία αυτή, παρουσιάζεται μια υλοποίηση σε FPGA ενός ψηφιακού κυκλώματος επεξεργασίας IP εικόνων. Το ψηφιακό κύκλωμα που υλοποιήθηκε αποτελεί μέρος ενός συστήματος συμπίεσης υψηλής απόδοσης που έχει

δημιουργηθεί σε λογισμικό [9]. Το σύστημα εκμεταλλεύεται τη μεγάλη ποσότητα πλεονάζουσας πληροφορίας που υπάρχει στις IP εικόνες.

Το κρίσιμότερο ζήτημα για τη σχεδίαση του ψηφιακού κυκλώματος ήταν η βελτιστοποίησή του ως προς την επιφάνεια κάλυψης, δεδομένου ότι αν ένα τελείως παράλληλο σύστημα επεξεργασίας υιοθετηθεί, πρέπει να χρησιμοποιηθούν περισσότερα από ένα FPGAs [15]. Στοχεύοντας στην ενσωμάτωση του κυκλώματος στο διαθέσιμο Virtex-E FPGA με επιφάνεια ισοδύναμη με 2 εκατομμύρια πύλες, καταφέραμε να δημιουργήσουμε ένα καινοτόμο σύστημα επεξεργασίας ικανό για εφαρμογές πραγματικού χρόνου, το οποίο διαδραματίζει το ρόλο επιταχυντή της συγκεκριμένης διεργασίας σε ένα ισχυρό σύστημα συμπίεσης τρισδιάστατων εικόνων.

#### Ευχαριστίες

Η εργασία πραγματοποιήθηκε στο πλαίσιο του προγράμματος «Πυθαγόρας», συγχρηματοδοτούμενο από την Ευρωπαϊκή Ένωση και το Υπουργείο Εθνικής Παιδείας και Θρησκευμάτων. Χρηματοδοτήθηκε μερικώς από τον Ειδικό Λογαριασμό Κονδυλίων Έρευνας (ΕΛΚΕ) του Εθνικού και Καποδιστριακού Πανεπιστημίου Αθηνών.

#### Αναφορές

- [1] Janusz Konrad “Visual Communications of Tomorrow: Natural, Efficient and Flexible”, *IEEE Communications Magazine*, vol. 39, no. 1, pp. 126-133, Jan. 2001
- [2] *Developers Handbook*, Stereographics 1997, [www.stereographics.com](http://www.stereographics.com)
- [3] M. Halle, “Autostereoscopic Displays and Computer Graphics”, *Computer Graphics*, 31(2), ACM SIGGRAPH, pp. 58-62, 1997
- [4] Yohsihiro Kajiki, Hiroshi Yoshikawa, and Toshio Honda, “Autostereoscopic 3-D Video Display Using Multiple Lightio Beams with Scanning”, *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 10, no. 2, pp 254-260, March 2000
- [5] Klaus Hopf, “An Autostereoscopic Display Providing Comfortable Viewing Conditions and a High Degree of Telepresence”, *IEEE Tans. on Circuits and Systems for Video Technology*, vol. 10, no. 3, pp 359-365, April 2000

- [6] H. Liao, S. Nakajima et. al., “Intra-operative Real-Time 3D Information Display System Based on Integral Videography”, *MICCAI' 01*, LNCS 2208, pp. 392-400, 2001
- [7] P. Harman, “Home Based 3D entertainment – an overview”, *Proc. ICIP(1)*, pp. 1-4, 2000
- [8] G. Lippman, “La Photographie Integrale”, *C. R. Acad. Sci*, 146, pp. 446-451, 1908
- [9] N. Sgouros, A. Andreou, M. Sangriotis, P. Papageorgas, D. Maroulis, N. Theofanous, “Compression of IP Images for Autostereoscopic 3D Imaging Applications”, *3rd International Symposium on Image and Signal Processing and Analysis (ISPA03)*, Rome, Italy, September 18-20, 2003
- [10] S. Rathnam, G. Slavenburg, “An Architectural Overview of the Programmable Multimedia Processor, TM-1”, *Proc. COMPCON '96*, pp. 319-326, 1996
- [11] S. Wong, S. Vassiliadis, S. Cotofana, “A Sum of Absolute Differences Implementation in FPGA Hardware”, *28<sup>th</sup> Euromicro Conference (EUROMICRO'02)*, pp. 183-186, Dortmund, Germany, September 04 - 06, 2002
- [12] R. Woods, A. Cassidy, J. Gray “VLSI Architectures for Field Programmable Gate Arrays: A Case Study” *IEEE Symposium on FPGAs for Custom Computing Machines*, 1996.
- [13] M. Martina, A. Molino, F. Vacca “Reconfigurable and Low Power 2D-DCT IP for Ubiquitous Multimedia Streaming” *2002 IEEE International Conference on Multimedia and Expo (ICME 2002)*, Vol. 2, August 26-29, pp. 177-180
- [14] Celoxica RC1000-PP development board: Hardware Reference, [www.celoxica.com](http://www.celoxica.com)
- [15] S. Wong, B. Stougie, S. Cotofana, “Alternatives in FPGA-based SAD Implementations”, *IEEE I.C. on Field Programmable Technology 2002 (FPT '02)*, Hong Kong, December 2002
- [16] Diogo Zandonai, Carro Luigi, Sergio Bampi, Altamiro Suzim “An Architecture for MPEG Motion Estimation”, *IWS'2001- VII Workshop IBERCHIP*, Montevideo, 2001. v.1. p.90-95