

ΕΞΑΓΟΝΤΑΣ ΤΗΝ ΑΝΤΙΛΗΠΤΗ ΠΟΙΟΤΗΤΑ ΒΙΝΤΕΟ ΑΠΟ ΜΕΤΡΗΣΕΙΣ ΤΗΣ ΜΕΘΟΔΟΥ VMAF ΣΕ ΕΠΙΠΕΔΟ ΠΛΑΙΣΙΟΥ

Η
ΜΕΤΑΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ ΕΞΕΙΔΙΚΕΥΣΗΣ

υποβάλλεται στην ορισθείσα
από την Γενική Συνέλευση Ειδικής Σύγκλησης
του Τμήματος Μηχανικών Η/Υ και Πληροφορικής
Εξεταστική Επιτροπή

από την

Σοφία Μπάτση

ως μέρος των υποχρεώσεων για τη απόκτηση του
ΜΕΤΑΠΤΥΧΙΑΚΟΥ ΔΙΠΛΩΜΑΤΟΣ ΣΤΗΝ ΠΛΗΡΟΦΟΡΙΚΗ
ΜΕ ΕΞΕΙΔΙΚΕΥΣΗ ΣΤΙΣ ΤΕΧΝΟΛΟΓΙΕΣ-ΕΦΑΡΜΟΓΕΣ

Πανεπιστήμιο Ιωαννίνων
Ιούνιος 2019

Εξεταστική Επιτροπή:

- **Λυσίμαχος Παύλος Κόντης**, Καθηγητής, Τμήμα Μηχανικών Η/Υ και Πληροφορικής, Πανεπιστήμιο Ιωαννίνων
- **Χριστόφορος Νίκου**, Καθηγητής, Τμήμα Μηχανικών Η/Υ και Πληροφορικής, Πανεπιστήμιο Ιωαννίνων
- **Κωνσταντίνος Παρσόπουλος**, Αναπλ. Καθηγητής, Τμήμα Μηχανικών Η/Υ και Πληροφορικής, Πανεπιστήμιο Ιωαννίνων

ΕΥΧΑΡΙΣΤΙΕΣ

Θα ήθελα να ευχαριστήσω θερμά τον καθηγητή κ. Λυσίμαχο Παύλο Κόντη για την υποστήριξη, την καθοδήγηση, τις προτάσεις και τις συμβουλές που μου παρείχε σε όλη τη διάρκεια εκπόνησης αυτής της εργασίας.

ΠΕΡΙΕΧΟΜΕΝΑ

ΕΥΧΑΡΙΣΤΙΕΣ	iv
ΠΕΡΙΕΧΟΜΕΝΑ	vi
ΚΑΤΑΛΟΓΟΣ ΠΙΝΑΚΩΝ	xi
ΠΕΡΙΛΗΨΗ	1
EXTENDED ABSTRACT	4
ΚΕΦΑΛΑΙΟ 1. ΕΙΣΑΓΩΓΗ	5
1.1. Θεματική περιοχή	5
1.2. Δομή της εργασίας	6
ΚΕΦΑΛΑΙΟ 2. ΜΕΤΑΒΑΣΗ ΑΠΟ ΤΗΝ ΠΟΙΟΤΗΤΑ ΥΠΗΡΕΣΙΑΣ ΣΤΗΝ ΠΟΙΟΤΗΤΑ ΕΜΠΕΙΡΙΑΣ	7
2.1. Ποιότητα υπηρεσίας	7
2.2. Ποιότητα Εμπειρίας	9
ΚΕΦΑΛΑΙΟ 3. ΜΕΤΡΙΚΕΣ ΠΟΙΟΤΗΤΑΣ ΒΙΝΤΕΟ	11
3.1. Η μετρική PSNR	12
3.2. Η μετρική SSIM	13
3.3. Η μετρική VQM	14
3.4. Η μετρική Video Multi-Method Assessment Fusion - VMAF	15
ΚΕΦΑΛΑΙΟ 4. ΒΕΛΤΙΩΝΟΝΤΑΣ ΤΟΝ ΑΛΓΟΡΙΘΜΟ VMAF	17
4.1. Μέθοδοι Χρονικής Συγκέντρωσης	18
4.1.1 Άθροισμα Minkowski και Αριθμητικός μέσος	18
4.1.2 Αρμονικός μέσος	18
4.1.3 Ρίζα Μέσου Τετραγώνου	19
4.1.4 Γεωμετρικός μέσος	19
4.1.5 Μέσος όρος των τελευταίων πλαισίων	20
4.1.6 K-th percentile	20
4.2. Βάσεις δεδομένων που χρησιμοποιήθηκαν	21
4.2.1 Η Βάση δεδομένων NETFLIX Video	21
4.2.2 Η Βάση δεδομένων Video Quality Expert Group HD3 (VQEG HD3)	23
ΚΕΦΑΛΑΙΟ 5. ΑΠΟΤΕΛΕΣΜΑΤΑ ΠΕΙΡΑΜΑΤΩΝ	24
5.1. Άθροισμα Minkowski	25
5.1.1 Αποτελέσματα με τη χρήση του NETFLIX Video Dataset	25
5.1.2 Αποτελέσματα με τη χρήση του VQEG HD3 Dataset	29
5.2. Μέσος όρος των τελευταίων πλαισίων	34
5.2.1. Αποτελέσματα με τη χρήση του NETFLIX Video Dataset	34
5.2.2. Αποτελέσματα με τη χρήση του VQEG HD3 Dataset	37

5.3. K-th percentile	40
5.3.1. Αποτελέσματα με τη χρήση του NETFLIX Video Dataset	40
5.3.2. Αποτελέσματα με τη χρήση του VQEG HD3 Dataset	43
5.4. Μέγιστη, ελάχιστη τιμή και διάμεσος	46
5.4.1. Αποτελέσματα με τη χρήση του NETFLIX Video Dataset	46
5.4.2. Αποτελέσματα με τη χρήση του VQEG HD3 Dataset	48
ΚΕΦΑΛΑΙΟ 6. ΣΥΜΠΕΡΑΣΜΑΤΑ	51
ΑΝΑΦΟΡΕΣ	53
ΠΑΡΑΡΤΗΜΑ	56

ΚΑΤΑΛΟΓΟΣ ΣΧΗΜΑΤΩΝ

ΕΙΚΟΝΑ 3-1 ΠΕΡΙΓΡΑΦΗ ΤΗΣ ΛΕΙΤΟΥΡΓΙΑΣ ΤΟΥ ΤΡΕΧΟΝΤΟΣ VMAF ΣΥΣΤΗΜΑΤΟΣ..	16
ΕΙΚΟΝΑ 5-1 ΓΡΑΦΗΜΑ ΑΠΟΤΕΛΕΣΜΑΤΩΝ VMAF SCORES ΜΕ ΜΕΘΟΔΟ ΧΡΟΝΙΚΗΣ ΣΥΓΚΕΝΤΡΩΣΗΣ ΤΟΝ ΑΡΙΘΜΗΤΙΚΟ ΜΕΣΟ (MINKOWSKI P=1).....	27
ΕΙΚΟΝΑ 5-2 ΓΡΑΦΗΜΑ ΑΠΟΤΕΛΕΣΜΑΤΩΝ VMAF SCORES ΜΕ ΜΕΘΟΔΟ ΧΡΟΝΙΚΗΣ ΣΥΓΚΕΝΤΡΩΣΗΣ ΤΟΝ ΑΡΜΟΝΙΚΟ ΜΕΣΟ (MINKOWSKI P=-1)	27
ΕΙΚΟΝΑ 5-3 ΓΡΑΦΗΜΑ ΑΠΟΤΕΛΕΣΜΑΤΩΝ VMAF SCORES ΜΕ ΜΕΘΟΔΟ ΧΡΟΝΙΚΗΣ ΣΥΓΚΕΝΤΡΩΣΗΣ ΤΟΝ ΓΕΩΜΕΤΡΙΚΟ ΜΕΣΟ (MINKOWSKI P→0).....	28
ΕΙΚΟΝΑ 5-4 ΓΡΑΦΗΜΑ ΑΠΟΤΕΛΕΣΜΑΤΩΝ VMAF SCORES ΜΕ ΜΕΘΟΔΟ ΧΡΟΝΙΚΗΣ ΣΥΓΚΕΝΤΡΩΣΗΣ ΤΗΝ ΡΙΖΑ ΜΕΣΟΥ ΤΕΤΡΑΓΩΝΟΥ (MINKOWSKI P=2)	28
ΕΙΚΟΝΑ 5-5 ΓΡΑΦΗΜΑ ΑΠΟΤΕΛΕΣΜΑΤΩΝ VMAF SCORES ΜΕ ΜΕΘΟΔΟ ΧΡΟΝΙΚΗΣ ΣΥΓΚΕΝΤΡΩΣΗΣ ΤΟ ΑΘΡΟΙΣΜΑ MINKOWSKI ΜΕ P=8.....	29
ΕΙΚΟΝΑ 5-6 ΓΡΑΦΗΜΑ ΑΠΟΤΕΛΕΣΜΑΤΩΝ VMAF SCORES ΜΕ ΜΕΘΟΔΟ ΧΡΟΝΙΚΗΣ ΣΥΓΚΕΝΤΡΩΣΗΣ ΤΟΝ ΑΡΙΘΜΗΤΙΚΟ ΜΕΣΟ (MINKOWSKI P=1)	31
ΕΙΚΟΝΑ 5-7 ΓΡΑΦΗΜΑ ΑΠΟΤΕΛΕΣΜΑΤΩΝ VMAF SCORES ΜΕ ΜΕΘΟΔΟ ΧΡΟΝΙΚΗΣ ΣΥΓΚΕΝΤΡΩΣΗΣ ΤΟΝ ΑΡΜΟΝΙΚΟ ΜΕΣΟ (MINKOWSKI P=-1)	31
ΕΙΚΟΝΑ 5-8 ΓΡΑΦΗΜΑ ΑΠΟΤΕΛΕΣΜΑΤΩΝ VMAF SCORES ΜΕ ΜΕΘΟΔΟ ΧΡΟΝΙΚΗΣ ΣΥΓΚΕΝΤΡΩΣΗΣ ΤΟΝ ΓΕΩΜΕΤΡΙΚΟ ΜΕΣΟ (MINKOWSKI P→0)	32
ΕΙΚΟΝΑ 5-9 ΓΡΑΦΗΜΑ ΑΠΟΤΕΛΕΣΜΑΤΩΝ VMAF SCORES ΜΕ ΜΕΘΟΔΟ ΧΡΟΝΙΚΗΣ ΣΥΓΚΕΝΤΡΩΣΗΣ ΤΗΝ ΡΙΖΑ ΜΕΣΟΥ ΤΕΤΡΑΓΩΝΟΥ (MINKOWSKI P=2)	32
ΕΙΚΟΝΑ 5-10 ΓΡΑΦΗΜΑ ΑΠΟΤΕΛΕΣΜΑΤΩΝ VMAF SCORES ΜΕ ΜΕΘΟΔΟ ΧΡΟΝΙΚΗΣ ΣΥΓΚΕΝΤΡΩΣΗΣ ΤΟ ΑΘΡΟΙΣΜΑ MINKOWSKI ΜΕ P=8.....	33
ΕΙΚΟΝΑ 5-11 ΓΡΑΦΗΜΑ ΑΠΟΤΕΛΕΣΜΑΤΩΝ VMAF SCORES ΜΕ ΜΕΘΟΔΟ ΧΡΟΝΙΚΗΣ ΣΥΓΚΕΝΤΡΩΣΗΣ ΤΟ ΑΘΡΟΙΣΜΑ MINKOWSKI ΜΕ P=10.....	33
ΕΙΚΟΝΑ 5-12 ΓΡΑΦΗΜΑ ΑΠΟΤΕΛΕΣΜΑΤΩΝ VMAF SCORES ΜΕ ΧΡΗΣΗ ΤΟΥ ΜΕΣΟΥ ΟΡΟΥ ΤΕΛΕΥΤΑΙΩΝ ΠΛΑΙΣΙΩΝ ΜΕ F=25.....	35
ΕΙΚΟΝΑ 5-13 ΓΡΑΦΗΜΑ ΑΠΟΤΕΛΕΣΜΑΤΩΝ VMAF SCORES ΜΕ ΧΡΗΣΗ ΤΟΥ ΜΕΣΟΥ ΟΡΟΥ ΤΕΛΕΥΤΑΙΩΝ ΠΛΑΙΣΙΩΝ ΜΕ F=50.....	35
ΕΙΚΟΝΑ 5-14 ΓΡΑΦΗΜΑ ΑΠΟΤΕΛΕΣΜΑΤΩΝ VMAF SCORES ΜΕ ΧΡΗΣΗ ΤΟΥ ΜΕΣΟΥ ΟΡΟΥ ΤΕΛΕΥΤΑΙΩΝ ΠΛΑΙΣΙΩΝ ΜΕ F=75.....	36
ΕΙΚΟΝΑ 5-15 ΓΡΑΦΗΜΑ ΑΠΟΤΕΛΕΣΜΑΤΩΝ VMAF SCORES ΜΕ ΧΡΗΣΗ ΤΟΥ ΜΕΣΟΥ ΟΡΟΥ ΤΕΛΕΥΤΑΙΩΝ ΠΛΑΙΣΙΩΝ ΜΕ F=100.....	36
ΕΙΚΟΝΑ 5-16 ΓΡΑΦΗΜΑ ΑΠΟΤΕΛΕΣΜΑΤΩΝ VMAF SCORES ΜΕ ΧΡΗΣΗ ΤΟΥ ΜΕΣΟΥ ΟΡΟΥ ΤΕΛΕΥΤΑΙΩΝ ΠΛΑΙΣΙΩΝ ΜΕ F=25.....	38
ΕΙΚΟΝΑ 5-17 ΓΡΑΦΗΜΑ ΑΠΟΤΕΛΕΣΜΑΤΩΝ VMAF SCORES ΜΕ ΧΡΗΣΗ ΤΟΥ ΜΕΣΟΥ ΟΡΟΥ ΤΕΛΕΥΤΑΙΩΝ ΠΛΑΙΣΙΩΝ ΜΕ F=50.....	38

ΕΙΚΟΝΑ 5-18 ΓΡΑΦΗΜΑ ΑΠΟΤΕΛΕΣΜΑΤΩΝ VMAF SCORES ΜΕ ΧΡΗΣΗ ΤΟΥ ΜΕΣΟΥ ΟΡΟΥ ΤΕΛΕΥΤΑΙΩΝ ΠΛΑΙΣΙΩΝ ΜΕ F=75.....	39
ΕΙΚΟΝΑ 5-19 ΓΡΑΦΗΜΑ ΑΠΟΤΕΛΕΣΜΑΤΩΝ VMAF SCORES ΜΕ ΧΡΗΣΗ ΤΟΥ ΜΕΣΟΥ ΟΡΟΥ ΤΕΛΕΥΤΑΙΩΝ ΠΛΑΙΣΙΩΝ ΜΕ F=100.....	39
ΕΙΚΟΝΑ 5-20 ΓΡΑΦΗΜΑ ΑΠΟΤΕΛΕΣΜΑΤΩΝ VMAF SCORES ΜΕ ΜΕΘΟΔΟ ΧΡΟΝΙΚΗΣ ΣΥΓΚΕΝΤΡΩΣΗΣ ΤΗΝ Κ-ΤΗ PERCENTILE ΜΕ K=5.....	41
ΕΙΚΟΝΑ 5-21 ΓΡΑΦΗΜΑ ΑΠΟΤΕΛΕΣΜΑΤΩΝ VMAF SCORES ΜΕ ΜΕΘΟΔΟ ΧΡΟΝΙΚΗΣ ΣΥΓΚΕΝΤΡΩΣΗΣ ΤΗΝ Κ-ΤΗ PERCENTILE ΜΕ K=10.....	41
ΕΙΚΟΝΑ 5-22 ΓΡΑΦΗΜΑ ΑΠΟΤΕΛΕΣΜΑΤΩΝ VMAF SCORES ΜΕ ΜΕΘΟΔΟ ΧΡΟΝΙΚΗΣ ΣΥΓΚΕΝΤΡΩΣΗΣ ΤΗΝ Κ-ΤΗ PERCENTILE ΜΕ K=20.....	42
ΕΙΚΟΝΑ 5-23 ΓΡΑΦΗΜΑ ΑΠΟΤΕΛΕΣΜΑΤΩΝ VMAF SCORES ΜΕ ΜΕΘΟΔΟ ΧΡΟΝΙΚΗΣ ΣΥΓΚΕΝΤΡΩΣΗΣ ΤΗΝ Κ-ΤΗ PERCENTILE ΜΕ K=25.....	42
ΕΙΚΟΝΑ 5-24 ΓΡΑΦΗΜΑ ΑΠΟΤΕΛΕΣΜΑΤΩΝ VMAF SCORES ΜΕ ΜΕΘΟΔΟ ΧΡΟΝΙΚΗΣ ΣΥΓΚΕΝΤΡΩΣΗΣ ΤΗΝ Κ-ΤΗ PERCENTILE ΜΕ K=5.....	44
ΕΙΚΟΝΑ 5-25 ΓΡΑΦΗΜΑ ΑΠΟΤΕΛΕΣΜΑΤΩΝ VMAF SCORES ΜΕ ΜΕΘΟΔΟ ΧΡΟΝΙΚΗΣ ΣΥΓΚΕΝΤΡΩΣΗΣ ΤΗΝ Κ-ΤΗ PERCENTILE ΜΕ K=10.....	44
ΕΙΚΟΝΑ 5-26 ΓΡΑΦΗΜΑ ΑΠΟΤΕΛΕΣΜΑΤΩΝ VMAF SCORES ΜΕ ΜΕΘΟΔΟ ΧΡΟΝΙΚΗΣ ΣΥΓΚΕΝΤΡΩΣΗΣ ΤΗΝ Κ-ΤΗ PERCENTILE ΜΕ K=20.....	45
ΕΙΚΟΝΑ 5-27 ΓΡΑΦΗΜΑ ΑΠΟΤΕΛΕΣΜΑΤΩΝ VMAF SCORES ΜΕ ΜΕΘΟΔΟ ΧΡΟΝΙΚΗΣ ΣΥΓΚΕΝΤΡΩΣΗΣ ΤΗΝ Κ-ΤΗ PERCENTILE ΜΕ K=25.....	45
ΕΙΚΟΝΑ 5-28 ΓΡΑΦΗΜΑ ΑΠΟΤΕΛΕΣΜΑΤΩΝ VMAF SCORES ΜΕ ΜΕΘΟΔΟ ΧΡΟΝΙΚΗΣ ΣΥΓΚΕΝΤΡΩΣΗΣ ΤΗΝ ΕΛΑΧΙΣΤΗ ΤΙΜΗ.....	47
ΕΙΚΟΝΑ 5-29 ΓΡΑΦΗΜΑ ΑΠΟΤΕΛΕΣΜΑΤΩΝ VMAF SCORES ΜΕ ΜΕΘΟΔΟ ΧΡΟΝΙΚΗΣ ΣΥΓΚΕΝΤΡΩΣΗΣ ΤΗ ΜΕΓΙΣΤΗ ΤΙΜΗ.	47
ΕΙΚΟΝΑ 5-30 ΓΡΑΦΗΜΑ ΑΠΟΤΕΛΕΣΜΑΤΩΝ VMAF SCORES ΜΕ ΜΕΘΟΔΟ ΧΡΟΝΙΚΗΣ ΣΥΓΚΕΝΤΡΩΣΗΣ ΤΗΝ ΔΙΑΜΕΣΟ.....	48
ΕΙΚΟΝΑ 5-31 ΓΡΑΦΗΜΑ ΑΠΟΤΕΛΕΣΜΑΤΩΝ VMAF SCORES ΜΕ ΜΕΘΟΔΟ ΧΡΟΝΙΚΗΣ ΣΥΓΚΕΝΤΡΩΣΗΣ ΤΗΝ ΕΛΑΧΙΣΤΗ ΤΙΜΗ.....	49
ΕΙΚΟΝΑ 5-32 ΓΡΑΦΗΜΑ ΑΠΟΤΕΛΕΣΜΑΤΩΝ VMAF SCORES ΜΕ ΜΕΘΟΔΟ ΧΡΟΝΙΚΗΣ ΣΥΓΚΕΝΤΡΩΣΗΣ ΤΗ ΜΕΓΙΣΤΗ ΤΙΜΗ.....	50
ΕΙΚΟΝΑ 5-33 ΓΡΑΦΗΜΑ ΑΠΟΤΕΛΕΣΜΑΤΩΝ VMAF SCORES ΜΕ ΜΕΘΟΔΟ ΧΡΟΝΙΚΗΣ ΣΥΓΚΕΝΤΡΩΣΗΣ ΤΗ ΔΙΑΜΕΣΟ.....	50

ΚΑΤΑΛΟΓΟΣ ΠΙΝΑΚΩΝ

ΠΙΝΑΚΑΣ 5-1 ΤΙΜΕΣ SRCC, PCC ΚΑΙ RMSE ΓΙΑ ΤΙΣ ΔΙΑΦΟΡΕΣ ΤΙΜΕΣ ΤΗΣ ΠΑΡΑΜΕΤΡΟΥ P.....	26
ΠΙΝΑΚΑΣ 5-2 ΤΙΜΕΣ SRCC, PCC ΚΑΙ RMSE ΓΙΑ ΤΙΣ ΔΙΑΦΟΡΕΣ ΤΙΜΕΣ ΤΗΣ ΠΑΡΑΜΕΤΡΟΥ P.....	30
ΠΙΝΑΚΑΣ 5-3 ΤΙΜΕΣ SRCC, PCC ΚΑΙ RMSE ΓΙΑ ΤΙΣ ΔΙΑΦΟΡΕΣ ΤΙΜΕΣ ΤΗΣ ΠΑΡΑΜΕΤΡΟΥ F.	34
ΠΙΝΑΚΑΣ 5-4 ΤΙΜΕΣ SRCC, PCC ΚΑΙ RMSE ΓΙΑ ΤΙΣ ΔΙΑΦΟΡΕΣ ΤΙΜΕΣ ΤΗΣ ΠΑΡΑΜΕΤΡΟΥ F.....	37
ΠΙΝΑΚΑΣ 5-5 ΤΙΜΕΣ SRCC, PCC ΚΑΙ RMSE ΓΙΑ ΤΙΣ ΔΙΑΦΟΡΕΣ ΤΙΜΕΣ ΤΗΣ ΠΑΡΑΜΕΤΡΟΥ Κ.....	40
ΠΙΝΑΚΑΣ 5-6 ΤΙΜΕΣ SRCC, PCC ΚΑΙ RMSE ΓΙΑ ΤΙΣ ΔΙΑΦΟΡΕΣ ΤΙΜΕΣ ΤΗΣ ΠΑΡΑΜΕΤΡΟΥ Κ.	43
ΠΙΝΑΚΑΣ 5-7 ΤΙΜΕΣ SRCC, PCC ΚΑΙ RMSE ΓΙΑ ΤΙΣ ΤΙΜΕΣ MIN, MAX ΚΑΙ MEDIAN....	46
ΠΙΝΑΚΑΣ 5-8 ΤΙΜΕΣ SRCC, PCC ΚΑΙ RMSE ΓΙΑ ΤΙΣ ΤΙΜΕΣ MIN, MAX ΚΑΙ MEDIAN....	48

ΠΕΡΙΛΗΨΗ

Μια από τις σημαντικότερες μετρικές οι οποίες χαρακτηρίζουν μια υπηρεσία πολυμέσων είναι η ποιότητα με την οποία την αντιλαμβάνεται ο τελικός χρήστης, γνωστή και ως ποιότητα εμπειρίας (Quality of Experience - QoE). Η μετρική αυτή σχετίζεται άμεσα με την επιλογή του χρήστη να συνεχίσει να χρησιμοποιεί (και ίσως να πληρώνει) την υπηρεσία ή να την διακόψει. Οι πάροχοι πολυμεσικών υπηρεσιών, προκειμένου να εξασφαλίσουν την ικανοποίηση των πελατών τους, αναπτύσσουν μεθόδους για τον αυτόματο (χωρίς την άμεση συμμετοχή του χρήστη) υπολογισμό της QoE.

Μία τέτοια μέθοδος είναι η Video Multi-Method Assessment Fusion (VMAF), που έχει προταθεί από τη Netflix. Η VMAF προσφέρει μια αυτοματοποιημένη εκτίμηση της ποιότητας εικόνας σε κάθε πλαίσιο μιας ακολουθίας βίντεο. Στη συνέχεια της διαδικασίας, υπολογίζεται ο αριθμητικός μέσος όρος των μετρήσεων ποιότητας ανά πλαίσιο, προκειμένου να ληφθεί μια εκτίμηση της συνολικής QoE της ακολουθίας βίντεο.

Σε αυτή την εργασία επικυρώνουμε την υπόθεση ότι ο αριθμητικός μέσος, ως μέθοδος χρονικής συγκέντρωσης (temporal pooling method), υποεκτιμά την συνεισφορά των πλαισίων κακής ποιότητας στον υπολογισμό της QoE, οδηγώντας σε τιμές υψηλότερες από αυτές που δίνονται από υποκειμενικές μετρήσεις. Δείχνουμε επίσης ότι το άθροισμα Minkowski (κατάλληλα παραμετροποιημένο) προσεγγίζει βέλτιστα τις υποκειμενικές μετρήσεις QoE. Η ποσοτικοποίηση αυτής της προσέγγισης αποτυπώνεται με τη χρήση των κυριότερων συντελεστών συσχέτισης, δηλαδή, Spearman Rank Correlation Coefficient (SRCC), Pearson Correlation Coefficient (PCC) καθώς και συμβατικά με την χρήση της ρίζας του μέσου τετραγωνικού σφάλματος, Root-Mean-Square-Error (RMSE).

Λέξεις κλειδιά— Αξιολόγηση ποιότητας βίντεο, χρονική συγκέντρωση, VMAF, Ποιότητα εμπειρίας.

EXTENDED ABSTRACT

One of the key elements that characterizes a multimedia service is its quality. This feature is an important factor that will lead a user to satisfaction and further use or avoidance. For this reason, it is especially necessary for multimedia service providers, in order to ensure the satisfaction of their customers, to develop metrics to calculate the user's QoE (Quality of Experience).

The Video Multimethod Assessment Fusion (VMAF) method, proposed by Netflix, offers an automated estimation of perceptual video quality for each frame of a video sequence. Then, the arithmetic mean of the per-frame quality measurements is taken by default, in order to obtain an estimate of the overall QoE of the video sequence. In this paper, we validate the hypothesis that the arithmetic mean conceals the bad quality frames, leading to an overestimation of the provided quality. We also show that the Minkowski mean (appropriately parametrized) approximates well the subjectively measured QoE, providing superior Spearman Rank Correlation Coefficient (SRCC), Pearson Correlation Coefficient (PCC), and Root-Mean-Square-Error (RMSE) scores.

Index Terms— Video quality assessment, pooling method, VMAF, QoE.

ΚΕΦΑΛΑΙΟ 1.

ΕΙΣΑΓΩΓΗ

1.1 Θεματική περιοχή

1.2 Δομή της εργασίας

1.1. Θεματική περιοχή

Η QoE είναι μετρική απόδοσης μίας υπηρεσίας η οποία επικεντρώνεται στον χρήστη/πελάτη, αποτυπώνοντας την εμπειρία και την ικανοποίησή του από την υπηρεσία που λαμβάνει. Μπορεί να μετρηθεί με τη συγκέντρωση αξιολογήσεων από τον χρήστη σε μια υποκειμενική δοκιμή αξιολόγησης ποιότητας. Σε αυτή την περίπτωση, η Βαθμολογία Μέσης Γνώμης (Mean Opinion Score - MOS) είναι μια ευρέως χρησιμοποιούμενη μέτρηση για την QoE. Οι βαθμολογίες MOS μπορούν επίσης να προβλεφθούν από αντικειμενικές μετρήσεις ποιότητας οι οποίες τυπικά έχουν αναπτυχθεί και εκπαιδευθεί χρησιμοποιώντας αξιολογήσεις του ανθρώπου. Η VMAF είναι μια μέθοδος μέτρησης ποιότητας βίντεο που αναπτύχθηκε από τη Netflix για την πρόβλεψη της υποκειμενικής ποιότητας βίντεο βάσει των MOS. Οι εκτιμήσεις ποιότητας ανά καρέ (frame) χρησιμοποιούνται για την παραγωγή συνολικής βαθμολογίας για ολόκληρο το βίντεο χρησιμοποιώντας μια μέθοδο χρονικής συγκέντρωσης (temporal pooling method). Η VMAF χρησιμοποιεί τον αριθμητικό μέσο ως προεπιλεγμένη μέθοδο χρονικής συγκέντρωσης, ώστε να καταλήξει σε ένα μοναδικό σκορ για ολόκληρο το βίντεο.

Τα αποτελέσματα στη βιβλιογραφία έχουν αποδείξει πως ο μέσος όρος ως μέθοδος χρονικής συγκέντρωσης “κρύβει” τα πλαίσια κακής ποιότητας, δίνοντας ίσα βάρη σε όλα τα πλαίσια, χωρίς να λαμβάνει υπόψη το επίπεδο παραμόρφωσής τους. Με βάση αυτήν την παρατήρηση, στην εργασία αυτή χρησιμοποιούνται διάφορες μέθοδοι χρονικής συγκέντρωσης πάνω σε διαφορετικές βάσεις δεδομένων, με στόχο την σύγκρισή τους με τον αριθμητικό μέσο, καταλήγοντας τελικά στην καλύτερη επιλογή της μεθόδου χρονικής συγκέντρωσης που θα μας δώσει συνολικό VMAF σκορ πιο κοντά στο MOS.

1.2. Δομή της εργασίας

Η δομή της εργασίας ορίζεται ως εξής. Στο Κεφάλαιο 2 περιγράφονται οι έννοιες της ποιότητας υπηρεσίας (QoS) και της ποιότητας εμπειρίας (QoE). Στο Κεφάλαιο 3 αναλύονται οι βασικότερες μετρικές ποιότητας βίντεο και περιγράφεται εκτενώς ο αλγόριθμος VMAF. Στο Κεφάλαιο 4 αποτυπώνονται οι μέθοδοι χρονικής συγκέντρωσης που χρησιμοποιούνται στην εργασία. Επίσης παρουσιάζονται τα σύνολα δεδομένων από βίντεο (video datasets) πάνω στα οποία εφαρμόστηκε η VMAF. Έπειτα στο Κεφάλαιο 5 παρουσιάζονται αναλυτικά τα αποτελέσματα χρήσης του αλγορίθμου VMAF στα δύο dataset με όλες τις μεθόδους χρονικής συγκέντρωσης που χρησιμοποιήθηκαν. Τέλος, στο Κεφάλαιο 6 παρατίθενται τα συμπεράσματα που αντλήθηκαν κατά την παρατήρηση των αποτελεσμάτων.

ΚΕΦΑΛΑΙΟ 2.

ΜΕΤΑΒΑΣΗ ΑΠΟ ΤΗΝ ΠΟΙΟΤΗΤΑ ΥΠΗΡΕΣΙΑΣ ΣΤΗΝ ΠΟΙΟΤΗΤΑ ΕΜΠΕΙΡΙΑΣ

2.1 Ποιότητα υπηρεσίας

2.2 Ποιότητα εμπειρίας

2.1. Ποιότητα υπηρεσίας

Ως ποιότητα υπηρεσίας ή Quality of Service (QoS) εννοούμε τη περιγραφή ή τη μέτρηση της συνολικής απόδοσης μιας υπηρεσίας, όπως ενός τηλεφωνικού δικτύου, ενός δικτύου υπολογιστών ή μιας υπηρεσίας video streaming, και ιδίως των επιδόσεων που βλέπουν οι χρήστες του δικτύου από κάποια υπηρεσία.

Ένας άλλος ορισμός ορίζει την ποιότητα υπηρεσίας ως την ικανότητα του δικτύου να παρέχει μία υπηρεσία σε ένα εγγυημένο επίπεδο ποιότητας. Περιλαμβάνει όλες τις λειτουργίες, τους μηχανισμούς και τις διαδικασίες ενός δικτύου, που εξασφαλίζουν την παροχή της συμφωνημένης ποιότητας υπηρεσίας μεταξύ του τελικού χρήστη και του πυρήνα του δικτύου. Πρόκειται για ένα ποσοτικό μέτρο της ποιότητας της υπηρεσίας, το οποίο συχνά μετριέται λόγω χαρακτηριστικών του δικτύου όπως απώλεια πακέτων, ρυθμός μετάδοσης bit, απόδοση, καθυστέρηση μετάδοσης, διαθεσιμότητα, jitter, κλπ.

Η ποιότητα της υπηρεσίας είναι σημαντική για εφαρμογές πολυμέσων με συνεχή ροή σε πραγματικό χρόνο, όπως η Voice over IP, video streaming και η IPTV, καθώς συχνά απαιτούν σταθερό ρυθμό μετάδοσης και είναι ευαίσθητα στην καθυστέρηση.

Στην περίπτωση του video streaming, η QoS σημαίνει ότι ο χρήστης θα πρέπει να παρακολουθεί ένα συγκεκριμένο περιεχόμενο, χωρίς διακοπή και χωρίς υποβάθμιση της ποιότητας του βίντεο που προκαλείται από τη διαδικασία παράδοσης.

Η QoS, όπως αναφέρεται παραπάνω, αφορά κυρίως την εκτίμηση τεχνικών χαρακτηριστικών ενός δικτύου, που σε συνδυασμό μπορούν να μας δώσουν μια εικόνα της απόδοσης της προσφερόμενης υπηρεσίας από το δίκτυο. Ωστόσο, πρέπει να ληφθεί υπ' όψη ότι ο υπολογισμός της QoS εξαρτάται από τα ιδιαίτερα χαρακτηριστικά της υπηρεσίας. Δηλαδή η QoS πρέπει να εκφράζεται με τις παραμέτρους του δικτύου. Για παράδειγμα διαφορετικές παράμετροι χαρακτηρίζουν μια απλή τηλεφωνία από μία υπηρεσία παροχής real-time video streaming, όπως αναφερθήκαμε παραπάνω.

Σύμφωνα και με την [1], οι παράμετροι δικτύου που συμβάλλουν στην εκτίμηση της QoS είναι:

Καθυστέρηση (Latency): Αυτή η παράμετρος είναι εγγενής σε όλα τα δίκτυα τηλεπικοινωνιών. Είναι ο χρόνος στον οποίο ένα πακέτο «φτάνει» από τον πάροχο στον τελικό χρήστη. Ο χρόνος αυτός είναι πολύ σημαντικός για τον καθορισμό της QoS, αφού μικροί χρόνοι καθυστέρησης είναι χαρακτηριστικό μιας «καλής ποιότητας» υπηρεσίας.

Διακύμανση καθυστέρησης (Jitter): Η διαφορά στους χρόνους καθυστέρησης μεταξύ διαδοχικών πακέτων μίας ροής δεδομένων. Οφείλεται σε εσωτερικές διαδικασίες δρομολόγησης του δικτύου και πιο συγκεκριμένα σε διαφορετικούς χρόνους αναμονής και επεξεργασίας στους δρομολογητές του δικτύου ή σε έκτακτες περιπτώσεις όπως η συμφόρηση ροής.

Απώλεια πακέτων (packet loss): Η αποτυχία μεταφοράς ενός ή περισσότερων πακέτων στον τελικό χρήστη του δικτύου.

Ρυθμαπόδοση (throughput): Ο χρονικός ρυθμός με τον οποίο μια οντότητα δικτύου λαμβάνει ή αποστέλλει δεδομένα. Όσο μεγαλύτερη ρυθμαπόδοση έχουμε, τόσο καλύτερη ποιότητα μπορεί να προσφέρει μια υπηρεσία.

Στη μετρική QoS είναι εξαιρετικά εύκολο να μετρηθεί η απόδοση του δικτύου της υπηρεσίας, δίνοντας ένα αριθμητικό μέτρο για αυτή. Όμως ύστερα από αυτό

τίθεται το ερώτημα στο χρήστη, εάν είναι ικανοποιημένος με την υπηρεσία που έλαβε. Επομένως παρατηρούμε, ότι αν και η μετρική της QoS είναι πολύ σημαντική, έχει ένα σημαντικό ελάττωμα: Αφορά αποκλειστικά την λειτουργία/απόδοση του δικτύου, η οποία είναι το πρώτο μισό μιας πολύ σημαντικής εξίσωσης.

Το άλλο μισό, στο οποίο εστιάζουν όλο και περισσότερο οι πάροχοι υπηρεσιών, είναι η Ποιότητα της Εμπειρίας (Quality of Experience – QoE).

2.2. Ποιότητα Εμπειρίας

Η QoE είναι μια μέτρηση απόδοσης που εστιάζει στον πελάτη για ένα συγκεκριμένο προϊόν ή υπηρεσία, σε αντίθεση με την QoS. Η QoE αφορά την εμπειρία που έχουν οι τελικοί χρήστες από μια υπηρεσία και την ικανοποίησή τους. Η QoE γίνεται ολοένα και πιο σημαντική καθώς το εμπόριο κερδίζει έδαφος στο Διαδίκτυο και η παροχή υπηρεσιών αρχίζει να πραγματοποιείται μέσω ενός υπολογιστή, εξασφαλίζοντας ότι ο πελάτης που βρίσκεται στο άλλο άκρο του υπολογιστή έχει μια θετική εμπειρία. Παρότι οι πάροχοι μιας διαδικτυακής υπηρεσίας επωφελούνται από την αυξημένη προσφορά χάρη στις ευκαιρίες που προσφέρει η γενικώς συνδεδεμένη γενιά μας, η απόσταση μεταξύ τους και του πελάτη δεν ήταν ποτέ μεγαλύτερη, με αποτέλεσμα η QoE να αποτελεί αναπόσπαστο μέρος κάθε νέας υπηρεσίας.

Λόγω της ποικιλίας των παραγόντων που μπορούν να επηρεάσουν την αλληλεπίδραση ενός ανθρώπου με μια υπηρεσία, αυτό που θεωρείται θετική εμπειρία για τους χρήστες είναι πολύ πιθανό να διαφέρει μεταξύ των υπηρεσιών. Κατά συνέπεια, το πρώτο βήμα μέτρησης της QoE μιας υπηρεσίας είναι να καθορίσει ποια είναι η επιθυμητή εμπειρία του πελάτη σε σχέση με την υπηρεσία.

Σε αντίθεση με την QoE, οι τελικοί χρήστες μιας υπηρεσίας δεν είναι το επίκεντρο της QoS, αλλά η απόδοση του δικτύου που τους παρέχεται. Οι χρήστες είναι φυσικά ένας παράγοντας, αλλά η QoS επικεντρώνεται περισσότερο στα τμήματα μεταξύ της IP και της εμπειρίας του χρήστη. Η QoS θεωρείται σήμερα ανεπαρκής για το λεπτομερή χαρακτηρισμό μιας παρεχόμενης υπηρεσίας, γιατί όπως αναφέρθηκε δεν ενσωματώνει οποιονδήποτε ανθρώπινο παράγοντα που επηρεάζει την

ποιότητα, ούτε αντικατοπτρίζει την πραγματική εμπειρία του χρήστη. Αυτό έχει ως αποτέλεσμα η ίδια η QoS, να μην μπορεί να εγγυηθεί το ίδιο επίπεδο ποιότητας για διαφορετικούς χρήστες.

Εκτός από τα τεχνικά χαρακτηριστικά του συστήματος άλλοι παράγοντες, όπως το πλαίσιο χρήσης, τα ανθρώπινα χαρακτηριστικά, το παραδοθέν περιεχόμενο, το κόστος της υπηρεσίας, κτλ., έχουν σημαντικό αντίκτυπο στην τελική αντίληψη για την εμπειρία του χρήστη και είναι πτυχές που δεν καταγράφει η QoS. Σε αυτό που αναφερθήκαμε παραπάνω «έρχεται» η QoE να συμπληρώσει την QoS και να καλύψει τους παραπάνω παράγοντες.

Σε μια εποχή όπου η ανάπτυξη των τηλεπικοινωνιών είναι μεγάλη και με τα δίκτυα 5ης γενιάς στον ορίζοντα, η QoE είναι σημαντική όσο ποτέ άλλοτε, διότι οι απαιτήσεις των χρηστών του δικτύου αυξάνονται ολοένα και περισσότερο. Λόγω της ραγδαίας ανάπτυξης των τηλεπικοινωνιών, παρατηρούμε αυξημένη δημοτικότητα σε υπηρεσίες βίντεο, όπως το video streaming, με συνέπεια να έχουμε μεγάλη ζήτηση από τους χρήστες του διαδικτύου, που χαρακτηρίζεται από υψηλή απαίτηση πόρων.

Όπως αναφέρθηκε παραπάνω η QoS και η QoE αποτελούν σημαντικά κριτήρια για δικτυακές εφαρμογές που σχετίζονται με video streaming. Μια τέτοια εταιρία που παρέχει συνεχόμενης προβολής ταινίες και σειρές δηλαδή video streaming είναι και η Netflix.

Η Netflix, ως εταιρία παροχής video streaming, διαχειρίζεται υπέρογκο πλήθος δεδομένων, το οποίο παρέχει στους πελάτες της. Γι' αυτό και προσπαθεί να δίνει και την καλύτερη εμπειρία στα μέλη της. Μια περιοχή που εστιάζει όσον αφορά την εμπειρία του χρήστη, είναι και η QoE, που αναφέρεται στην εμπειρία του χρήστη από τη στιγμή που θα πατήσει το κουμπί play στην εφαρμογή της Netflix. Η Netflix έχει δεσμευτεί στα μέλη της για την παροχή εξαιρετικών υπηρεσιών streaming και γι' αυτό έχει δημιουργήσει αποκλειστικά μια ομάδα που ειδικεύεται στο να βελτιώσει την ποιότητα εμπειρίας των υπηρεσιών της. Εξετάζει αρκετές μετρικές, για να προσδιορίσει ποιες απ' αυτές επηρεάζουν την ικανοποίηση του χρήστη και τελικά την ποιότητα εμπειρίας.

ΚΕΦΑΛΑΙΟ 3.

ΜΕΤΡΙΚΕΣ ΠΟΙΟΤΗΤΑΣ ΒΙΝΤΕΟ

3.1 Η μετρική Peak Signal to Noise Ratio - PSNR

3.2 Η μετρική Structural Similarity Image Metric - SSIM

3.3 Η μετρική Video Quality Monitor - VQM

3.4 Η μετρική Video Multi-Method Assessment Fusion - VMAF

Η μέτρηση ποιότητας βίντεο με διάφορες μετρικές εξασφαλίζει πως οι φορείς μπορούν να μεταφέρουν στους χρήστες την επιθυμητή QoE. Τα τελευταία χρόνια, ο ερευνητής Staelens [2], δήλωσε ότι η υποκειμενική εκτίμηση ποιότητας βίντεο με βάση το MOS είναι η πιο ακριβής μέθοδος ώστε να διατηρηθεί η ανθρώπινη κρίση σε σχέση με την ποιότητα του βίντεο. Παρόλα αυτά, οι υποκειμενικές αξιολογήσεις είναι εξωφρενικά χρονοβόρες και δεν μπορούν να εξάγουν αποτελέσματα σε πραγματικό χρόνο.

Για αυτόν τον λόγο η βιομηχανία έχει αναζητήσει μετρικές ποιότητας βίντεο που είναι εύκολο να υπολογιστούν και ανακλούν την υποκειμενική ανθρώπινη εκτίμηση της ποιότητας του βίντεο. Για τις περισσότερες από τις προτεινόμενες αντικειμενικές μετρήσεις ποιότητας βίντεο που υπάρχουν στη βιβλιογραφία, το σήμα αναφοράς (αρχικό βίντεο ή βίντεο αναφοράς) είναι πλήρως διαθέσιμο. Αυτές οι μετρικές αναφέρονται ως Full Reference (FR) μετρικές ποιότητας βίντεο. Αξίζει να σημειωθεί εδώ ότι πρακτικά, συχνά το βίντεο αναφοράς είναι μη προσπελάσιμο. Στην περίπτωση αυτή χρησιμοποιούνται οι μετρικές που ονομάζονται No Reference (NR) μετρικές, ενώ υπάρχει και ένα τρίτο είδος μετρικών ποιότητας βίντεο κατά το οποίο το βίντεο αναφοράς δεν είναι πλήρως διαθέσιμο και αντί για αυτό εξάγονται

ορισμένα χαρακτηριστικά από αυτό και μεταδίδονται στο σύστημα αξιολόγησης ποιότητας ως πλευρικές πληροφορίες που θα βοηθήσουν να αξιολογηθεί η ποιότητα του βίντεο. Αυτές οι μετρικές αναφέρονται ως Reduced-Reference (RR) μετρικές.

Μερικές από τις ευρέως χρησιμοποιούμενες μετρικές αναλύονται στη συνέχεια του κεφαλαίου.

3.1. Η μετρική PSNR

Η Peak Signal to Noise Ratio (PSNR) είναι μια παραδοσιακή FR μετρική ποιότητας σήματος, που μετράτε σε decibel. Προέρχεται απευθείας από το μέσο τετραγωνικό σφάλμα (Mean Square Error - MSE) ή την τετραγωνική του ρίζα (Root Mean Square Error - RMSE). Το σφάλμα υπολογίζεται σε όλα τα εικονοστοιχεία (pixels) του βίντεο σε σχέση με το αντίστοιχο βίντεο αναφοράς. Για μια ακολουθία βίντεο, το μέσο τετραγωνικό σφάλμα μεταξύ κάθε ζεύγους i αντίστοιχων πλαισίων αναφοράς και τον παραμορφωμένων βίντεο είναι:

$$MSE = \frac{1}{N} \sum_{i=1}^N (x_i - y_i)^2 \quad (3.1)$$

όπου N είναι ο αριθμός των pixels της εικόνας ή του βίντεο και x_i και y_i είναι τα i -οστά pixels στην εικόνα (βίντεο) αναφοράς και στην παραμορφωμένη αντίστοιχα.

Το PSNR για κάθε ζεύγος i τελικά ορίζεται ως:

$$PSNR = 10 \log_{10} \frac{L^2}{MSE} \quad (3.2)$$

όπου L είναι η μέγιστη δυνατή τιμή των pixel.

Η δημοτικότητά του ως μέτρο ποιότητας βίντεο προέρχεται κυρίως από το γεγονός ότι είναι γρήγορο στον υπολογισμό και μαθηματικά απλό στην κατανόηση.

Το PSNR παρέχει με επιτυχία μια αριθμητική τιμή κατά τη σύγκριση ενός βίντεο αναφοράς με το κωδικοποιημένο βίντεο αυτού, ωστόσο αρκετές μελέτες που σχετίζονται με βίντεο χαμηλής ανάλυσης [3][4][5][6][7][8][9], έχουν συγκρίνει τη πρόβλεψη ποιότητας βίντεο PSNR με υποκειμενικά δεδομένα και έχουν καταλήξει

ότι μια τέτοια απλή αντικειμενική μέτρηση σχετίζεται ελάχιστα με την ανθρώπινη αντίληψη της ποιότητας.

3.2. Η μετρική SSIM

Η Structural Similarity Image Metric (SSIM) είναι μια FR μέτρηση ποιότητας εικόνας, που εισήχθη το 2004 από τον ερευνητή Wang [10][11]. Η αρχική έκδοση του μοντέλου δημιουργήθηκε στο Laboratory for Image and Video Engineering (LIVE) στο Πανεπιστήμιο του Τέξας στο Austin. Δημιουργήθηκε για να βελτιώσει τις παραδοσιακές μετρικές ποιότητας βίντεο όπως η PSNR και το MSE, για αυτόν τον λόγο ενώ οι πρώτες βασίζονται στα απόλυτα σφάλματα, η SSIM είναι βασισμένη στο μοντέλο HVS (Human visual system).

Υπολογίζει ένα σκορ για κάθε ένα μεμονωμένο pixel, χρησιμοποιώντας ένα παράθυρο από γειτονικά pixel. Αυτά τα σκορ υπολογίζονται κατά μέσο όρο ώστε να παράγουν μια συνολική βαθμολογία για ολόκληρη την εικόνα, σε σχέση πάντα με την εικόνα αναφοράς. Ορίζεται ως εξής:

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (3.3)$$

όπου x και y είναι τα δύο παράθυρα με κοινό μέγεθος $N \times N$, μ_x και μ_y είναι ο μέσος όρος των x και y αντίστοιχα, σ_x^2 και σ_y^2 η διακύμανση των x και y αντιστοίχως, σ_{xy} η συνδιακύμανση των x και y και τέλος $c_1 = (k_1L)^2$ και $c_2 = (k_2L)^2$ δύο μεταβλητές για την σταθεροποίηση της διαίρεσης με αδύναμους παρονομαστές, όπου L είναι το δυναμικό εύρος των τιμών των εικονοστοιχείων και $k_1 = 0.01$ και $k_2 = 0.03$ σταθερές.

Η μετρική αυτή έχει υιοθετηθεί και για την εύρεση ποιότητας ολόκληρου βίντεο.

3.3. Η μετρική VQM

Η Video Quality Monitor (VQM) προσφέρεται ως εμπορικό εργαλείο το οποίο δεν απαιτεί ουσιαστική αναφορά σε αρχικό βίντεο. Αποτελεί μια Reduced-Reference μετρική. Μπορεί να μετρήσει αντικειμενικά το πόσο θολό είναι το βίντεο και το frame rate.

Το μοντέλο VQM [12] είναι μια μετρική ποιότητας βίντεο που βασίζεται σε Διακριτό Μετασχηματισμό Συνημιτόνου (Discrete Cosine Transform- DCT). Αυτή η μέτρηση βασίζεται στο μοντέλο απλοποιημένης ανθρώπινης χωρικής και χρονικής ευαισθησίας αντίθεσης. Υπολογίζει τη παραμόρφωση ενός συμπιεσμένου βίντεο σε 4 βήματα:

- Για κάθε πλαίσιο το μοντέλο εκτελεί DCT μετασχηματισμό για 8x8 pixel μπλοκ από τα αρχικά πλαίσια του βίντεο και για τα μπλοκ του παραμορφωμένου πλαισίου βίντεο.
- Το μοντέλο μετατρέπει τους συντελεστές DCT σε τιμές τοπικής αντίθεσης με τη χρήση συνιστωσών DC του κάθε μπλοκ, για το αρχικό και το συμπιεσμένο βίντεο.
- Στη συνέχεια τους μετατρέπει σε συντελεστές διαφοράς, χρησιμοποιώντας τη στατική και δυναμική χωρική Συνάρτηση Ευαισθησίας Αντίθεσης (Contrast Sensitivity Function - CFS).
- Τέλος, οι συντελεστές του προηγούμενου βήματος, από την αρχική και την συμπιεσμένη ακολουθία, αφαιρούνται με σκοπό να παραχθούν τιμές διαφοράς. Το μοντέλο, ενσωματώνει τη μάσκα αντίθεσης με την βοήθεια της συνάρτησης \max και μετά χρησιμοποιεί ως μέθοδο χρονικής συγκέντρωσης τον αριθμητικό μέσο όρο.

Η βαθμολογία VQM μειώνεται καθώς η ποιότητα του συμπιεσμένου βίντεο αυξάνεται, και είναι μηδενική όταν το συμπιεσμένο βίντεο δεν έχει καμία απώλεια.

Το εργαλείο VQM είναι διαθέσιμο στο διαδίκτυο για ερευνητικούς σκοπούς [13] .

3.4. Η μετρική Video Multi-Method Assessment Fusion - VMAF

Σύμφωνα με τα [14] και [15], προκειμένου να πετύχει εξαιρετική ποιότητα του βίντεο που παρέχει στα μέλη της, η Netflix δημιούργησε την μέθοδο Video Multimethod Assessment Fusion - VMAF που υπολογίζει την ποιότητα του video streaming και είναι έντονα συσχετισμένη με τις υποκειμενικές βαθμολογίες (MOS).

Πρόκειται για μια πλήρους αναφοράς μετρική η οποία έχει ως κύριο στόχο της να προσεγγίσει όσο το δυνατόν καλύτερα την ανθρώπινη αντιληπτή ποιότητα βίντεο. Η VMAF εξάγει μια σειρά βασικών μετρικών ποιότητας βίντεο ως χαρακτηριστικά και τα τροφοδοτεί σε ένα Support Vector Machine (SVM) [16][17]. Αυτό επιτρέπει στη VMAF να μην αλλοιώσει την βαρύτητα του κάθε ξεχωριστού χαρακτηριστικού από τα οποία αποτελείται.

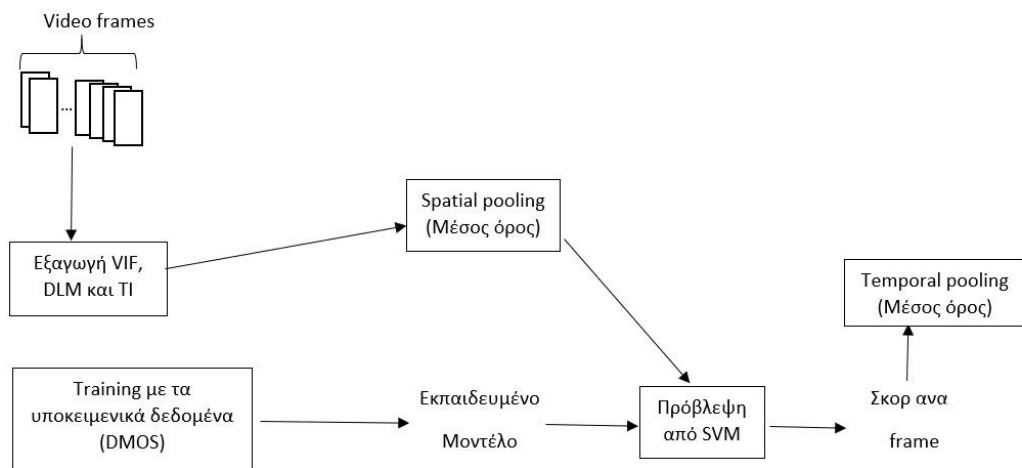
Το σύστημα VMAF αποτελείται από τα ακόλουθα βήματα (Εικόνα 3.1) : εξαγωγή χαρακτηριστικών, training/testing και χρονική συγκέντρωση (temporal pooling) των τελικών σκορ.

Στο πρώτο βήμα εξάγεται μια σειρά μετρικών ως αντιληπτά χαρακτηριστικά:

- **Detail Loss Metric (DLM)** [18]. Η DLM καταγράφει τις απώλειες λεπτομέρειας και τις ξεχωρίζει από τις προσθετικές βλάβες με την δημιουργία ενός αλγορίθμου wavelet και υπολογίζει ένα σταθμισμένο άθροισμα των DLM τιμών πάνω από τέσσερις διαφορετικές κλίμακες. Οι απώλειες λεπτομέρειας αναφέρονται στην απώλεια χρήσιμων οπτικών λεπτομερειών που επηρεάζουν την προβολή του περιεχομένου και οι προσθετικές βλάβες αντιπροσωπεύουν τις πλεονάζουσες πληροφορίες των οποίων η εμφάνισή τους αποσπά δυσάρεστα την προσοχή του θεατή.
- **Visual Information Fidelity VIF** [19]. Η VIF είναι FR μετρική που βασίζεται στα στατιστικά της φυσικής σκηνής και στην έννοια των πληροφοριών εικόνας που εξάγονται από το HVS. Δημιουργήθηκε από τους Hamid R Sheikh και Alan Bovik στο Πανεπιστήμιο του Τέξας στο Όστιν το 2006 και έδειξε ότι συσχετίζεται πολύ καλά με την ανθρώπινη κρίση της οπτικής ποιότητας. Καταγράφει τις απώλειες οπτικής ποιότητας πληροφοριών και υπολογίζεται σε τέσσερις κλίμακες, αποδίδοντας τέσσερα χαρακτηριστικά VIF.

- Temporal Information- TI** (οι διαφορές φωτεινότητας μεταξύ των ζευγών πλαισίων. Το χαρακτηριστικό TI (η χρονική αντιληπτή παρατήρηση) βασίζεται στη διαφορά κίνησης που είναι η διαφορά μεταξύ των τιμών των εικονοστοιχείων (pixel) στην ίδια θέση αλλά σε διαδοχικούς χρόνους ή πλαισίων ως συνάρτηση του χρόνου. Ουσιαστικά έχει ως στόχο να καταγράψει χρονικές επιδράσεις που οφείλονται σε μεταβολές κινήσεων που προσδιορίζονται ποσοτικά από τις διαφορές φωτεινότητας.

Όλα τα παραπάνω, καταλήγουν σε συνολικά έξι χαρακτηριστικά ανά πλαίσιο, από τα οποία υπολογίζεται ο αριθμητικός μέσος όρος, ανά χαρακτηριστικό για ολόκληρο το βίντεο. Αυτό, τελικά καταλήγει, σε έξι χαρακτηριστικά ανά βίντεο. Στο δεύτερο βήμα (κατά την διαδικασία training), αυτές οι έξι τιμές των χαρακτηριστικών, μαζί με το αντίστοιχο υποκειμενικό σκορ (MOS), εισάγονται σε ένα μοντέλο SVR. Κατά το testing, η VMAF προβλέπει μια τιμή ανά πλαίσιο και στο τελευταίο βήμα υπολογίζει τον αριθμητικό μέσο όρο τους, ώστε να προβλέψει τη συνολική ποιότητα του βίντεο, το VMAF σκορ, δηλαδή, για ολόκληρο το βίντεο.



Εικόνα 3-1 Περιγραφή της λειτουργίας του τρέχοντος VMAF συστήματος

ΚΕΦΑΛΑΙΟ 4.

ΒΕΛΤΙΩΝΟΝΤΑΣ ΤΟΝ ΑΛΓΟΡΙΘΜΟ VMAF

4.1 Μέθοδοι Χρονικής Συγκέντρωσης

4.2 Βάσεις δεδομένων που χρησιμοποιήθηκαν

Η αξιολόγηση ποιότητας βίντεο ασχολείται με την πρόβλεψη της αντιληπτής ποιότητάς του. Ένα από τα σημαντικότερα σημεία αυτής της πρόβλεψης είναι η χρονική συγκέντρωση (temporal pooling) των σκορ ανά πλαίσιο με σκοπό την εξαγωγή ενός τελικού σκορ για όλο το βίντεο. Τα πλαίσια κακής ποιότητας επηρεάζουν σημαντικά την συνολική υποκειμενική αντίληψη της ποιότητας του βίντεο [20][21]. Αυτό, δηλώνει την ανάγκη για την εύρεση κατάλληλων μεθόδων συγκέντρωσης που εξάγουν συνολικό σκορ δίνοντας έμφαση στα πλαίσια κακής ποιότητας.

Ο πλέον διαδεδομένος -λόγω της απλότητάς του- τρόπος συγκέντρωσης των βαθμολογιών των πλαισίων, είναι ο υπολογισμός του αριθμητικού μέσου όρου για την πρόβλεψη της συνολικής ποιότητας. Πολλές μετρικές ποιότητας βίντεο τον έχουν χρησιμοποιήσει και μια από αυτές είναι και η VMAF όπως αναφέρθηκε στο προηγούμενο κεφάλαιο.

Ωστόσο, η χρήση του αριθμητικού μέσου όρου ως μέθοδος χρονικής συγκέντρωσης, μπορεί να μην είναι συνεπής σε σχέση με τον τρόπο με τον οποίο ο άνθρωπος παρατηρεί και αξιολογεί την ποιότητα ενός βίντεο [22].

Για την εξάλειψη του παραπάνω προβλήματος, δοκιμάστηκαν κάποιες γνωστές μέθοδοι, που αναλύονται στην συνέχεια του κεφαλαίου, οι οποίες λαμβάνουν υπόψη τους τα πλαίσια κακής ποιότητας, με σκοπό την εύρεση αυτής η οποία μας δίνει σκορ πιο κοντά στα υποκειμενικά σκορ. Για να επικυρωθούν τα αποτελέσματα,

εφαρμόσαμε τις μεθόδους χρονικής συγκέντρωσης σε δύο από τις ευρέως γνωστές βάσεις δεδομένων, όπου και αυτές θα αναλυθούν παρακάτω.

4.1. Μέθοδοι Χρονικής Συγκέντρωσης

4.1.1 Άθροισμα Minkowski και Αριθμητικός μέσος

Ο αριθμητικός μέσος όρος, δίνει ίσα βάρη σε όλα τα πλαίσια, με αποτέλεσμα να τα αντιμετωπίζει όλα το ίδιο, σε αντίθεση με το άθροισμα Minkowski όπου το τελικό σκορ επηρεάζεται από τα πλαίσια που έχουν υποστεί μεγαλύτερη αλλοίωση.

Έστω t το εκάστοτε πλαίσιο, $OM(t)$ είναι η τιμή VMAF για το t πλαίσιο, T είναι ο συνολικός αριθμός των πλαισίων και p είναι ο εκθέτης Minkowski. Το άθροισμα Minkowski δίνεται από τον τύπο:

$$OM_{Mink} = \left[\frac{1}{T} \sum_{t=1}^T OM^p(t) \right]^{1/p}, \quad p \neq 0 \quad (4.1)$$

Όσο μεγαλύτερη είναι η τιμή της παραμέτρου Minkowski, p , τόσο περισσότερο δίνεται έμφαση στα πλαίσια με την υψηλότερη ποιότητα.

Ωστόσο, για συγκεκριμένες τιμές της παραμέτρου p ο τύπος της (4.1) μετασχηματίζεται σε άλλες γνωστές μεθόδους χρονικής συγκέντρωσης. Όπως για παράδειγμα για $p=1$, ο τύπος μετασχηματίζεται στον γνωστό τύπο αριθμητικού μέσου:

$$OM_{mean} = \sum_{t=1}^T OM(t) \quad (4.2)$$

4.1.2 Αρμονικός μέσος

Ο αρμονικός μέσος ενός συνόλου ορίζεται ως το αντίστροφο του αριθμητικού μέσου των αντίστροφων παρατηρήσεων, δηλαδή ο τύπος του αθροίσματος Minkowski (4.1) με $p = -1$:

$$OM_{Harmonic} = \left[\frac{1}{T} \sum_{t=1}^T OM^{-1}(t) \right]^{-1} \quad (4.3)$$

Ο αρμονικός μέσος όρος συχνά παράγει αποτελέσματα πολύ κοντά στον αριθμητικό μέσο όρο, εκτός από τις περιπτώσεις που στο δείγμα μας έχουμε τιμές outliers. Τότε, δίνει έμφαση στην επίδραση των πλαισίων με μικρότερες τιμές.

4.1.3 Ρίζα Μέσου Τετραγώνου

Στην στατιστική, η ρίζα μέσου τετραγώνου (Root mean square), γνωστή και ως Quadratic Mean, ορίζεται ως ο αριθμητικός μέσος όρος των τετραγώνων ενός συνόλου τιμών. Αν στον τύπο του αθροίσματος Minkowski (4.1), όπου $p=2$ τότε προκύπτει ο τύπος της ρίζας μέσου τετραγώνου:

$$OM_{RMS} = \left[\frac{1}{T} \sum_{t=1}^T OM^2(t) \right]^{1/2} \quad (4.4)$$

4.1.4 Γεωμετρικός μέσος

Ο Γεωμετρικός μέσος είναι μια τιμή όπου δείχνει την κεντρική τάση ή αλλιώς την τυπική τιμή ενός συνόλου αριθμών. Ορίζεται ως η n -οστή ρίζα του γινομένου n αριθμών. Δηλαδή για ένα σύνολο αριθμών x_1, x_2, \dots, x_n , ο γεωμετρικός μέσος όρος ορίζεται ως εξής:

$$\left(\prod_{i=1}^n x_i \right)^{\frac{1}{n}} = \sqrt[n]{x_1 x_2 \dots x_n} \quad (4.5)$$

Ένα σημαντικό χαρακτηριστικό του γεωμετρικού όρου σε σύγκριση με τον αριθμητικό μέσο είναι ότι ο πρώτος επηρεάζεται λιγότερο από παρατηρήσεις με πολύ μεγάλες τιμές.

Ωστόσο δεν υπάρχει συγκεκριμένο ρ στην Minkowski (4.1) , ώστε να αναπαριστά τον γεωμετρικό μέσο, διότι επιτυγχάνεται μέσω του ορίου της παραμέτρου ρ να τείνει στο 0. Για τιμές επομένως της παραμέτρου ρ που τείνουν στο 0, παίρνουμε τον Γεωμετρικό μέσο.

4.1.5 Μέσος όρος των τελευταίων πλαισίων

Ο μέσος όρος των τελευταίων πλαισίων (Mean last frames) στηρίζεται στο φαινόμενο recency effect (το φαινόμενο κατά το οποίο ο άνθρωπος είναι πιθανόν να θυμάται καλύτερα τα πιο πρόσφατα παρουσιαζόμενα σε αυτόν στοιχεία ή εμπειρίες).

Έτσι, αυτό που κάνει η (4.6) είναι να υπολογίζει τον αριθμητικό μέσο όρο των τελευταίων F πλαισίων.

$$OM_{meanF} = \frac{1}{F} \sum_{t=T-F}^T OM(t) \quad (4.6)$$

Όσο πιο υψηλές είναι οι τιμές του F τόσο πιο πολύ εξασθενεί το φαινόμενο recency effect.

4.1.6 K -th percentile

Ως k -th percentile ενός ταξινομημένου συνόλου ορίζεται το $k\%$ των χαμηλότερων τιμών του συνόλου αυτού. Δεδομένου ενός συνόλου, τα στοιχεία πρώτα ταξινομούνται με αύξουσα τάξη μεγέθους (στην περίπτωση μας τα πλαίσια ταξινομούνται από το πλαίσιο με το μικρότερο σκορ προς το πλαίσιο με το μεγαλύτερο) με τις χαμηλότερες τιμές να δηλώνουν το k -th percentile του συνόλου αυτού.

Η παραπάνω μέθοδος χρονικής συγκέντρωσης βασίζεται στην υπόθεση πως τα κακής ποιότητας πλαίσια μπορούν να κυριαρχήσουν και να επηρεάσουν την συνολική υποκειμενική αντίληψη της ποιότητας ενός βίντεο.

4.2. Βάσεις δεδομένων που χρησιμοποιήθηκαν

Οι βάσεις δεδομένων αξιολόγησης της ποιότητας επιτρέπουν στους ερευνητές να αξιολογούν την απόδοση των αλγορίθμων αξιολόγησης ποιότητας και να συμβάλλουν στην επίτευξη του απώτερου στόχου της αντικειμενικής έρευνας αξιολόγησης ποιότητας που ταιριάζει με την ανθρώπινη αντίληψη.

Δεδομένου ότι οι προτεινόμενες μέθοδοι χρονικής συγκέντρωσης δεν αποδίδουν το ίδιο σε όλες τις βάσεις δεδομένων, για να αξιολογηθούν και να συγκριθούν με βάση την απόδοσή τους, πρέπει να δοκιμαστούν και να παρουσιαστούν πειράματα σε ένα ευρύ φάσμα συνόλου δεδομένων (datasets).

Δοκιμάστηκαν σε δύο από τα βασικότερα διαθέσιμα στο κοινό σύνολα δεδομένων ποιότητας εικόνων (image quality datasets), το NETFLIX Video Dataset [14] και το Video Quality Expert Group HD3 (VQEG HD3) [23].

4.2.1 Η Βάση δεδομένων NETFLIX Video

Στην Netflix, για να αξιολογήσουν και να συγκρίνουν τον αλγόριθμο αξιολόγησης ποιότητας βίντεο VMAF, δημιούργησαν ένα σύνολο δεδομένων που σχετίζεται με την περίπτωση χρήσης τους, και αυτό γιατί παρόλο που υπάρχουν πολλές διαθέσιμες βάσεις δεδομένων για τον σχεδιασμό και τη δοκιμή μετρήσεων ποιότητας βίντεο, καμία δεν διαθέτει την ποικιλομορφία του περιεχομένου βίντεο που παρουσιάζει η Netflix.

Η υπηρεσία streaming της Netflix παρουσιάζει πολλές προκλήσεις, όπως για παράδειγμα:

- **Τα χαρακτηριστικά του βίντεο:** Η Netflix φέρει μια τεράστια συλλογή από ταινίες και εκπομπές, με ποικίλο περιεχόμενο όπως animation, ντοκιμαντέρ, κινούμενα σχέδια, κινηματογραφικές ταινίες κ.ά., τα οποία παρουσιάζουν επίσης ποικίλα χαρακτηριστικά όπως σκούρες σκηνές ή πολύ φωτεινά χρώματα, θόρυβο κ.ά.
- **Πηγή των αντικειμένων:** Η υπηρεσία streaming της Netflix μεταφέρεται με τη χρήση του πρωτοκόλλου ελέγχου μετάδοσης (TCP), επομένως οι μόνοι

τύποι αλλοίωσης της ποιότητας εικόνας είναι κατά την κωδικοποίηση: τα αντικείμενα συμπίεσης (compression artifacts) και τα αντικείμενα κλιμάκωσης (scaling artifacts).

Για να δημιουργήσουν ένα dataset προσαρμοσμένο στις ανάγκες της Netflix, επιλέχτηκε ένα δείγμα από 34 αρχικά clips (βίντεο αναφοράς), διάρκειας 6 δευτερολέπτων το καθένα, βγαλμένα από τον κατάλογο ταινιών και προγραμμάτων της Netflix ώστε να καλύψουν ένα ευρύ φάσμα χαρακτηριστικών υψηλού επιπέδου όπως κίνηση κάμερας, κοντινό πρόσωπο, κινούμενα σχέδια, νερό αλλά και χαρακτηριστικά χαμηλού επιπέδου όπως θόρυβος, πλούσιο χρώμα, υφή.

Χρησιμοποιώντας τα βίντεο αναφοράς, κωδικοποίησαν τα βίντεο με H.264 / AVC σε αναλύσεις που κυμαίνονται από 384 x 288 έως 1920 x 1080 και bitrate από 375 kbps έως 20000 kbps, με αποτέλεσμα να παραχθούν περίπου 300 παραμορφωμένα βίντεο.

Στη συνέχεια εκτέλεσαν υποκειμενικά τεστ για να προσδιορίσουν με ποιόν τρόπο οι απλοί παρατηρητές βαθμολογούν τις παραμορφώσεις των κωδικοποιημένων βίντεο. Οι παρατηρητές κάθισαν σε περιβάλλον ειδικά διαμορφωμένο να μοιάζει με σαλόνι σπιτιού και κλήθηκαν να αξιολογήσουν τη παραμόρφωση των βίντεο σε κλίμακα από 1 (πολύ ενοχλητική) έως 5 (δεν είναι αισθητή). Τα βίντεο παρουσιάστηκαν στους παρατηρητές σε αντιπαράθεση με το βίντεο αναφοράς.

Οι βαθμολογίες συγκεντρώθηκαν και συνδυάστηκαν για να δημιουργήσουν ένα Differential Mean Opinion Score – DMOS [24] (είναι σύνηθες οι βαθμολογίες να αντισταθμίζονται σε σχέση με τη βαθμολογία του βίντεο αναφοράς) για κάθε παραμορφωμένο βίντεο και κανονικοποιήθηκαν στο εύρος 0 έως 100. Το σύνολο των βίντεο αναφοράς, των παραμορφωμένων βίντεο και των βαθμολογιών τους αποτελούν το NETFLIX Video Dataset.

Στην παρούσα εργασία χρησιμοποιήθηκε ένα μέρος του NETFLIX Video Dataset, το οποίο είναι διαθέσιμο στο ευρύ κοινό.

4.2.2 Η Βάση δεδομένων Video Quality Expert Group HD3 (VQEG HD3)

Η ομάδα VQEG αρχικά βασίστηκε στη βασική υποκειμενική μεθοδολογία και την αντικειμενική ανάπτυξη και επαλήθευση εργαλείων για την αξιολόγηση ποιότητας του βίντεο. Αρχικά εστίασε στην μέτρηση της εικόνας βίντεο, και κινήθηκε στην ποιότητα της εμπειρίας (QoE). Ένα από τα projects που δημιούργησαν είναι και το HDTV, το οποίο αποτελείται από πέντε σύνολα δεδομένων (datasets), vqeghd1, vqeghd2, vqeghd3, vqeghd4, vqeghd5, vqeghd6. Το τεστ HDTV απευθύνεται σε τέσσερις μορφές βίντεο (1080p στα 25 και 29,97 frames-per-second, και 1080i στα 50 και 59.94 fields-per-second) και δύο μορφές βίντεο (720p στα 50 και 59.94 frames-per-second). Αυτή η δοκιμή HDTV αφορούσε τρεις τύπους μοντέλων: πλήρης αναφορά (FR), μειωμένη αναφορά (RR), και καμία αναφορά (NR). Τα μοντέλα FR έχουν πλήρη πρόσβαση στο βίντεο πηγής. Στο δικό μας πείραμα, χρησιμοποιούμε μόνο το vqeghd3, και μόνο τα βίντεο από SRC01 έως SRC09 (βίντεο αναφοράς) και HRC04, HRC07 και από HRC16 έως HRC21 (παραγόμενα βίντεο) τα οποία δημιουργήθηκαν με κωδικοποίηση MPEG-2 και H.264. Στην παρούσα εργασία χρησιμοποιήθηκε ένα μέρος της παραπάνω βάσης, αυτό που προτείνεται από τους δημιουργούς της VMAF στον διαδίκτυο.

ΚΕΦΑΛΑΙΟ 5.

ΑΠΟΤΕΛΕΣΜΑΤΑ ΠΕΙΡΑΜΑΤΩΝ

5.1 Άθροισμα Minkowski

5.2 Μέσος όρος τελευταίων πλαισίων

5.3 k-th percentile

5.4 Μέγιστη, ελάχιστη τιμή και διάμεσος

Όπως αναφέρθηκε στο κεφάλαιο 3, οι περισσότερες μέθοδοι χρονικής συγκέντρωσης που χρησιμοποιήθηκαν, απαιτούν παραμέτρους. Για αυτόν τον λόγο όλοι οι παραμετρικοί αλγόριθμοι χρονικής συγκέντρωσης δοκιμάστηκαν για πολλές τιμές των παραμέτρων, με σκοπό να βρεθεί η βέλτιστη τιμή που μεγιστοποιεί τη συσχέτιση μεταξύ του VMAF score και των υποκειμενικών score.

Για να επικυρώσουμε τα αποτελέσματα, χρησιμοποιήσαμε τρία κριτήρια βελτιστοποίησης: τον συντελεστή συσχέτισης Pearson (Pearson Correlation Coefficient - PCC) , τον συντελεστή συσχέτισης Spearman Rank (Spearman Rank Correlation Coefficient - SRCC) και τη ρίζα μέσου τετραγωνικού σφάλματος (Root Mean Square Error - RMSE).

Ο συντελεστής PCC είναι μέτρο της γραμμικής συσχέτισης μεταξύ των αντικειμενικών προβλέψεων και των υποκειμενικών βαθμολογιών. Παίρνει τιμές μεταξύ +1 και -1, με την τιμή 1 να σημαίνει τον πλήρη γραμμικό συσχετισμό των δύο μεταβλητών.

Ο συντελεστής SRCC μετράει την σχέση μονοτονίας μεταξύ των αντικειμενικών προβλέψεων και των υποκειμενικών βαθμολογιών, με τιμές μεταξύ +1 και -1, με την τιμή 1 να σημαίνει την πλήρη μονοτονική συσχέτιση των δύο μεταβλητών.

Τέλος, ο συντελεστής RMSE είναι μια πολύ συχνά χρησιμοποιούμενη μέτρηση της τυπικής απόκλισης των σφαλμάτων πρόβλεψης. Τα σφάλματα πρόβλεψης είναι ένα μέτρο για το πόσο μακριά βρίσκονται τα σημεία δεδομένων της παλινδρόμησης.

5.1. Άθροισμα Minkowski

Όπως προαναφέρθηκε στην αρχή του κεφαλαίου αυτού, για το άθροισμα Minkowski ως μέθοδος χρονικής συγκέντρωσης, χρησιμοποιήθηκαν διάφορες τιμές για την Minkowski παράμετρο p , ώστε να επιλεγεί εκείνη που εξάγει το επιθυμητό αποτέλεσμα. Ως επιθυμητό αποτέλεσμα χαρακτηρίζονται οι τιμές που πλησιάζουν τις τιμές MOS του εκάστοτε εξεταζόμενου dataset. Ως τιμές της Minkowski παραμέτρου p δοκιμάστηκαν οι εξής: -1, 0.5, 1, 2, 2.5, 3, 3.5, 4, 5, 8, 10, 50, 100.

5.1.1 Αποτελέσματα με τη χρήση του NETFLIX Video Dataset

Στον Πίνακα 5.1 φαίνονται οι τιμές των SRCC, PCC και RMSE για τις διάφορες τιμές της παραμέτρου p . Τα καλύτερα αποτελέσματα προκύπτουν όταν η Minkowski παράμετρος p πάρει την τιμή 8. Ξεκινώντας από την τιμή -1 και ανεβαίνοντας μέχρι την τιμή 8, η συσχέτιση μεταξύ του VMAF σκορ και του MOS είναι ισχυρότερη. Από την τιμή 8 και μετά οι συσχέτιση μεταξύ των δύο σκορ ελαττώνεται.

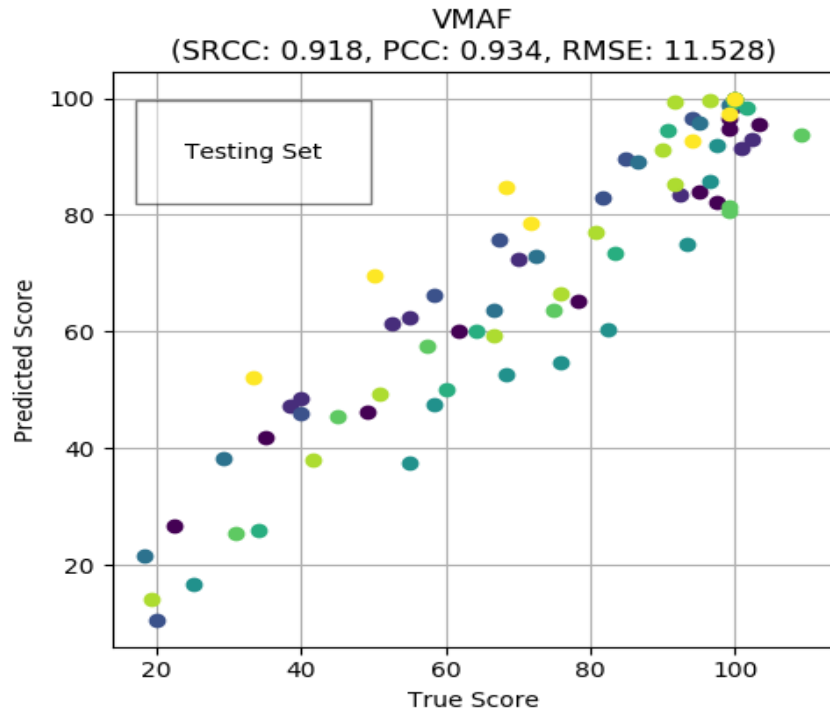
Η Εικόνα 5.1 δείχνει το γράφημα όπου στον άξονα X έχει τις βαθμολογίες DMOS και στον άξονα Y το VMAF σκορ χρησιμοποιώντας ως μέθοδο χρονικής συγκέντρωσης τον Αριθμητικό μέσο, προκαθορισμένο από την VMAF (Minkowski με $p=1$). Η Εικόνα 5.2 δείχνει τα αποτελέσματα με την χρήση της μεθόδου χρονικής συγκέντρωσης Αρμονικού μέσου (Minkowski με $p=-1$), η Εικόνα 5.3 δείχνει τα αποτελέσματα με την χρήση της μεθόδου χρονικής συγκέντρωσης Γεωμετρικού μέσου (Minkowski με $p \rightarrow 0$), η Εικόνα 5.4 δείχνει τα αποτελέσματα με την χρήση της μεθόδου χρονικής

συγκέντρωσης Ρίζα μέσου τετραγώνου (Minkowski με $p=2$) και τέλος η Εικόνα 5.5 δείχνει το γράφημα των αποτελεσμάτων με την χρήση της μεθόδου χρονικής συγκέντρωσης αθροίσματος Minkowski με $p=8$.

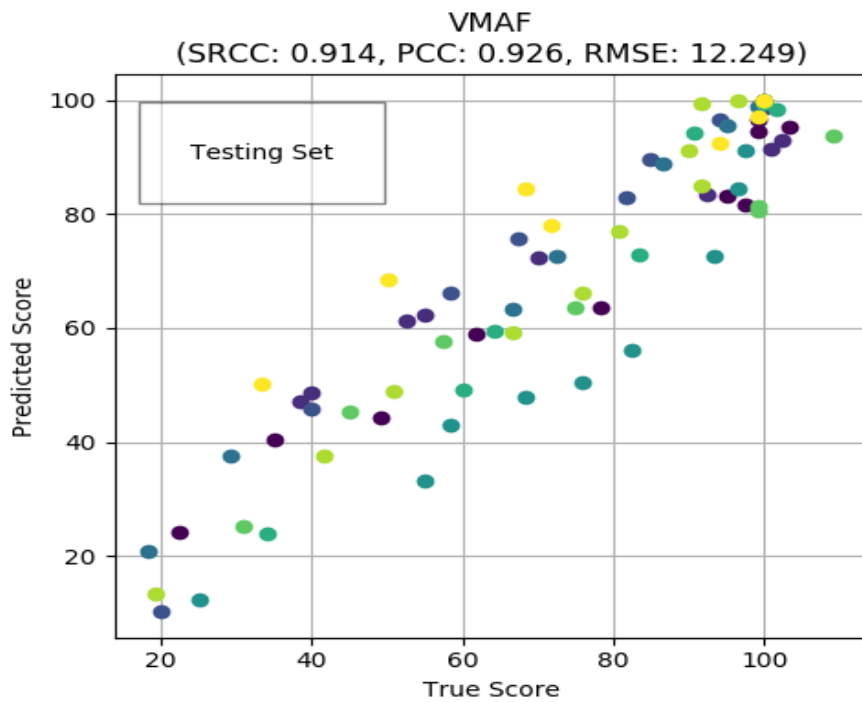
Κάθε τελεία διαφορετικού χρώματος αναπαριστά ένα παραμορφωμένο βίντεο. Οι τελείες ίδιου χρώματος υποδηλώνουν βίντεο που προέρχονται από το ίδιο βίντεο αναφοράς. Τέλος, εξαιτίας της μεταβλητότητας και της κανονικοποίησης των βίντεο αναφοράς στην κλίμακα 0-100, μερικά αποτελέσματα DMOS μπορεί να υπερβαίνουν τα 100.

Πίνακας 5-1 Τιμές SRCC, PCC και RMSE για τις διάφορες τιμές της παραμέτρου p .

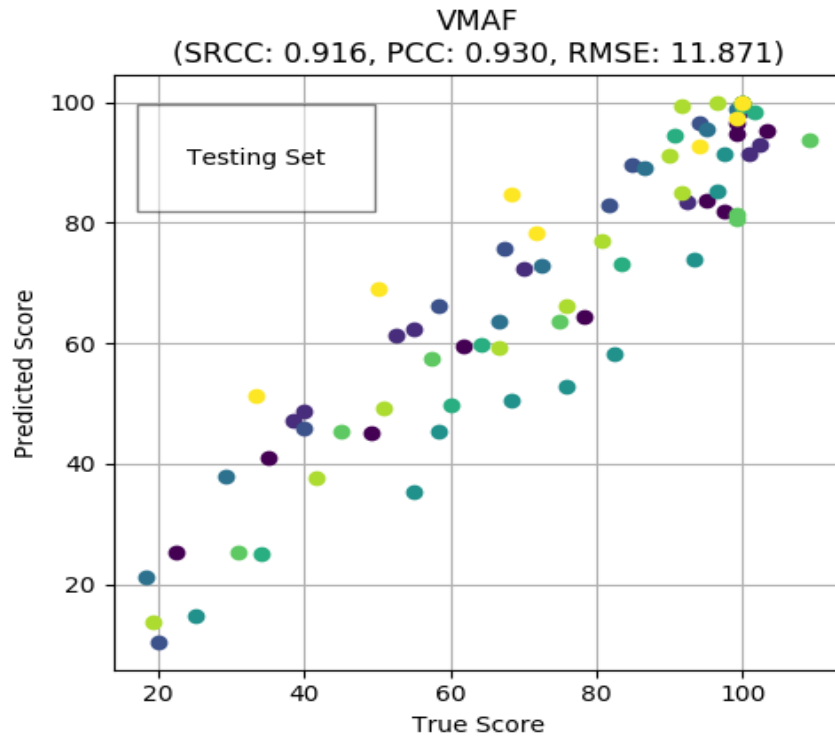
VMAF				
Μέθοδοι χρονικής συγκέντρωσης		SRCC	PCC	RMSE
Άθροισμα Minkowski	$p = -1$ (Αρμονικός μέσος)	0.914	0.926	12.249
	$p \rightarrow 0$ (Γεωμετρικός μέσος)	0.916	0.930	11.871
	$p = 0.5$	0.918	0.934	11.529
	$p = 1$ (VMAF default)	0.918	0.934	11.529
	$p = 2$ (Ρίζα μέσου τετραγώνου)	0.92	0.938	11.234
	$p = 2.5$	0.921	0.939	11.106
	$p = 3$	0.921	0.94	10.99
	$p = 3.5$	0.922	0.941	10.886
	$p = 4$	0.922	0.942	10.793
	$p = 5$	0.923	0.944	10.642
	$p = 8$	0.924	0.946	10.422
	$p = 10$	0.921	0.945	10.434
	$p = 50$	0.907	0.924	12
	$p = 100$	0.895	0.913	12.741



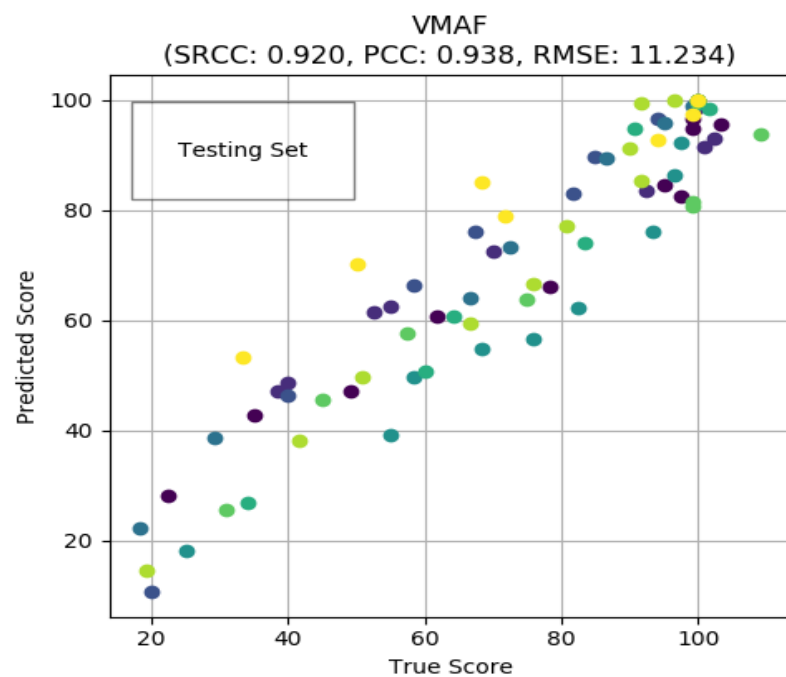
Εικόνα 5-1 Γράφημα αποτελεσμάτων VMAF scores με μέθοδο χρονικής συγκέντρωσης τον Αριθμητικό μέσο (Minkowski $p=1$)



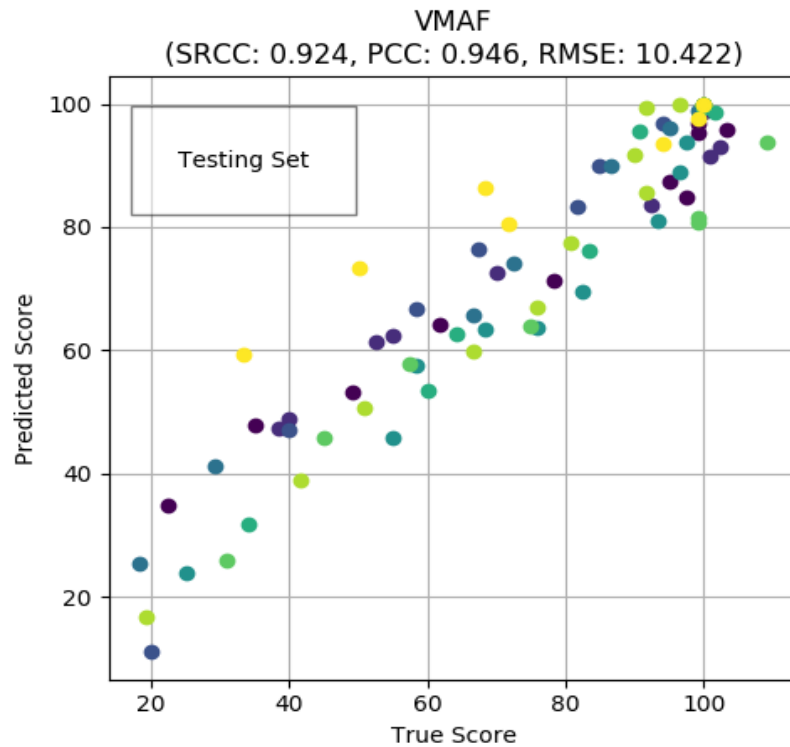
Εικόνα 5-2 Γράφημα αποτελεσμάτων VMAF scores με μέθοδο χρονικής συγκέντρωσης τον Αρμονικό μέσο (Minkowski $p=-1$)



Εικόνα 5-3 Γράφημα αποτελεσμάτων VMAF scores με μέθοδο χρονικής συγκέντρωσης τον Γεωμετρικό μέσο (Minkowski $p \rightarrow 0$)



Εικόνα 5-4 Γράφημα αποτελεσμάτων VMAF scores με μέθοδο χρονικής συγκέντρωσης την Ρίζα μέσου τετραγώνου (Minkowski $p=2$)



Εικόνα 5-5 Γράφημα αποτελεσμάτων VMAF scores με μέθοδο χρονικής συγκέντρωσης το άθροισμα Minkowski με $p=8$

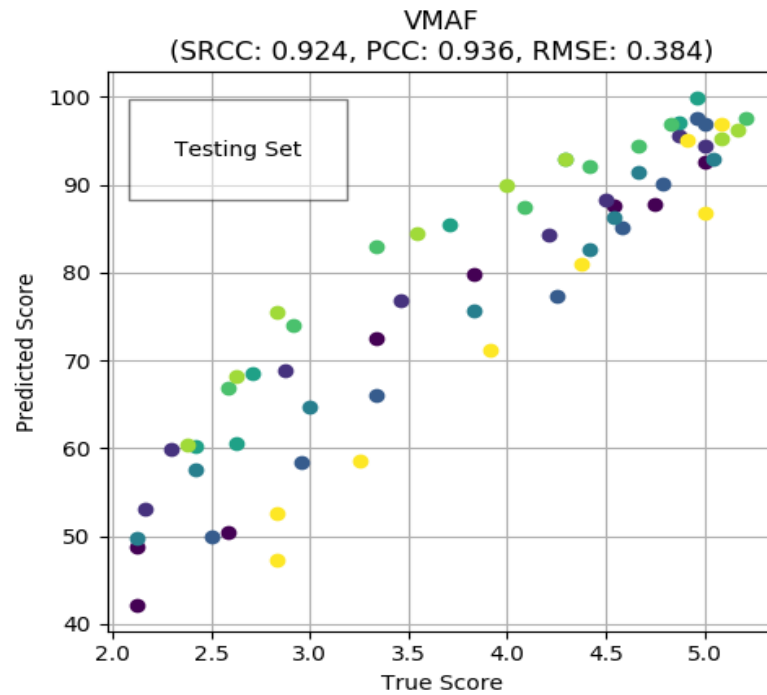
5.1.2 Αποτελέσματα με τη χρήση του VQEG HD3 Dataset

Στον Πίνακα 5.2 φαίνονται οι τιμές των SRCC, PCC και RMSE για τις διάφορες τιμές της παραμέτρου p χρησιμοποιώντας ως μέθοδο χρονικής συγκέντρωσης το άθροισμα Minkowski στο VQEG HD3 Dataset. Τα καλύτερα αποτελέσματα προκύπτουν όταν η Minkowski παράμετρος p πάρει την τιμή 8 και την τιμή 10.

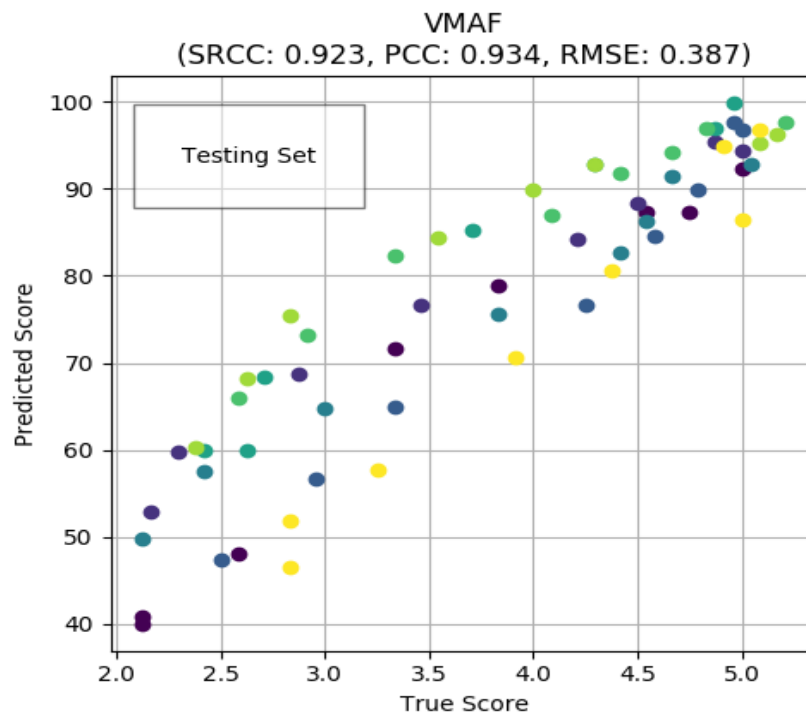
Πίνακας 5-2 Τιμές SRCC, PCC και RMSE για τις διάφορες τιμές της παραμέτρου p .

VMAF				
Μέθοδοι χρονικής συγκέντρωσης		SRCC	PCC	RMSE
Άθροισμα Minkowski	$p = -1$ (Αρμονικός μέσος)	0.923	0.934	0.387
	$p \rightarrow 0$ (Γεωμετρικός μέσος)	0.924	0.935	0.386
	$p = 0.5$	0.925	0.935	0.385
	$p = 1$ (VMAF default)	0.924	0.936	0.384
	$p = 2$ (Ρίζα μέσου τετραγώνου)	0.923	0.936	0.383
	$p = 2.5$	0.923	0.937	0.382
	$p = 3$	0.922	0.937	0.382
	$p = 3.5$	0.923	0.937	0.381
	$p = 4$	0.925	0.938	0.381
	$p = 5$	0.925	0.938	0.38
	$p = 8$	0.926	0.939	0.377
	$p = 10$	0.926	0.939	0.376
	$p = 50$	0.906	0.922	0.419
	$p = 100$	0.891	0.909	0.453

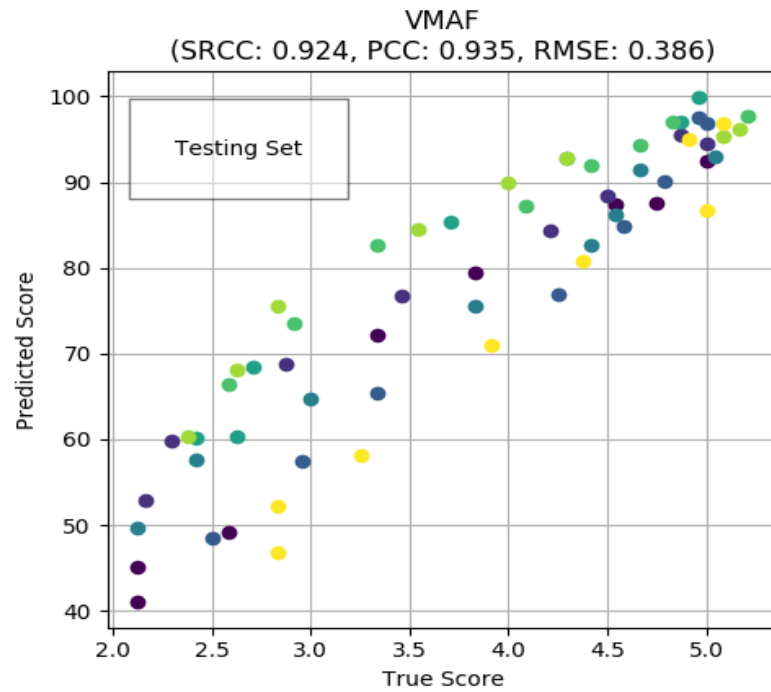
Οι Εικόνες 5.6, 5.7, 5.8, 5.9, 5.10, 5.11 δείχνουν τα αποτελέσματα με τη χρήση του αριθμητικού μέσου όρου (Mean), του Αρμονικού μέσου, του Γεωμετρικού μέσου, της Ρίζα μέσου τετραγώνου, του αθροίσματος Minkowski με $p=8$ και του αθροίσματος Minkowski με $p=10$ ως μέθοδος χρονικής συγκέντρωσης, αντίστοιχα.



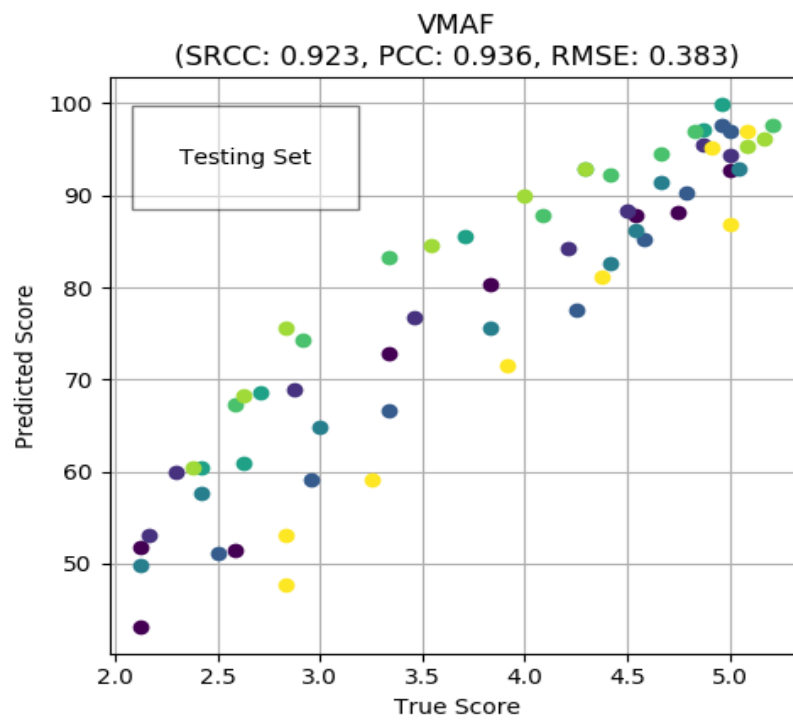
Εικόνα 5-6 Γράφημα αποτελεσμάτων VMAF scores με μέθοδο χρονικής συγκέντρωσης τον Αριθμητικό μέσο (Minkowski $p=1$)



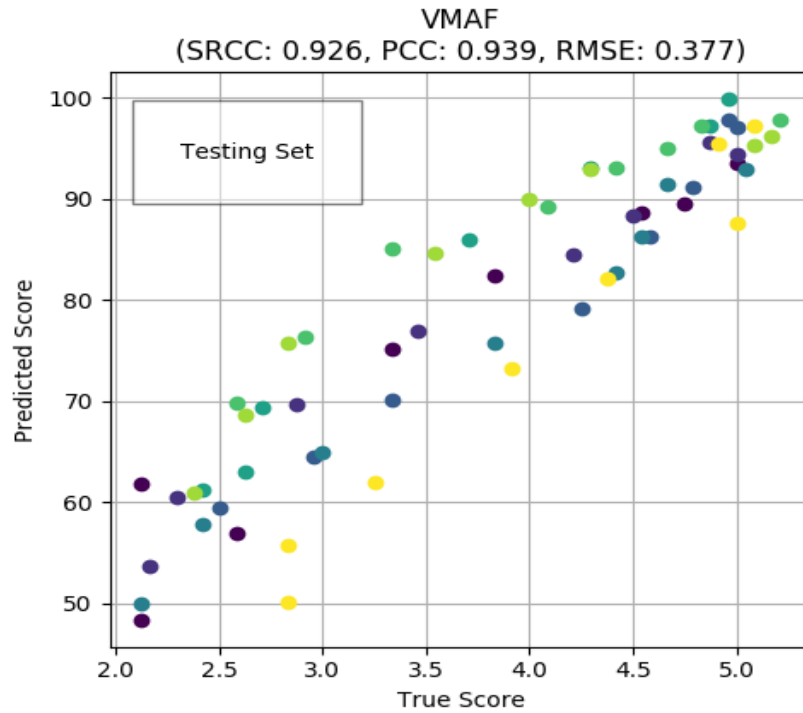
Εικόνα 5-7 Γράφημα αποτελεσμάτων VMAF scores με μέθοδο χρονικής συγκέντρωσης τον Αρμονικό μέσο (Minkowski $p=-1$)



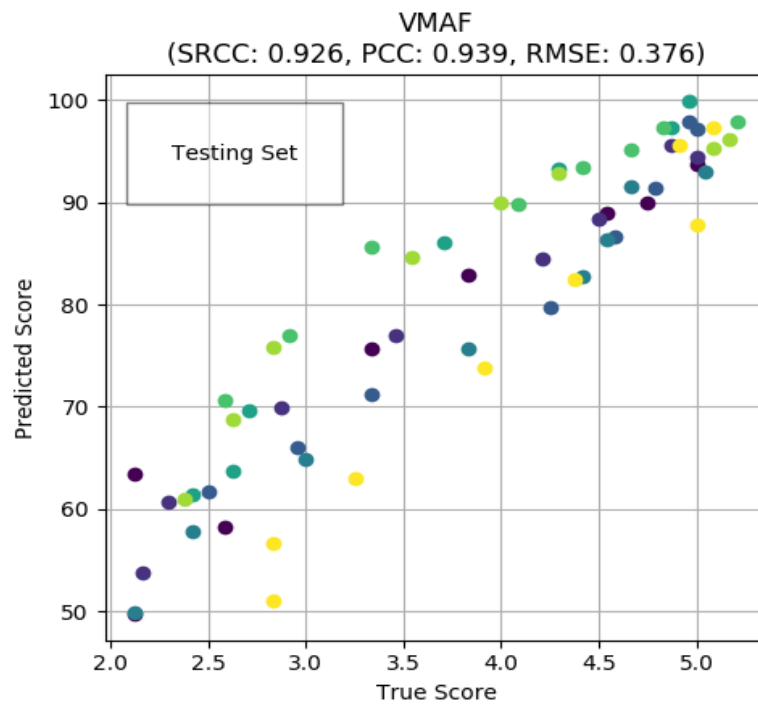
Εικόνα 5-8 Γράφημα αποτελεσμάτων VMAF scores με μέθοδο χρονικής συγκέντρωσης τον Γεωμετρικό μέσο (Minkowski $p \rightarrow 0$)



Εικόνα 5-9 Γράφημα αποτελεσμάτων VMAF scores με μέθοδο χρονικής συγκέντρωσης την Ρίζα μέσου τετραγώνου (Minkowski $p=2$)



Εικόνα 5-10 Γράφημα αποτελεσμάτων VMAF scores με μέθοδο χρονικής συγκέντρωσης το άθροισμα Minkowski με $p=8$



Εικόνα 5-11 Γράφημα αποτελεσμάτων VMAF scores με μέθοδο χρονικής συγκέντρωσης το άθροισμα Minkowski με $p=10$

5.2. Μέσος όρος των τελευταίων πλαισίων

Για τη χρήση της μεθόδου χρονικής συγκέντρωσης Μέσου όρου τελευταίων πλαισίων, δοκιμάσαμε διάφορες τιμές για την παράμετρο F της (4.6) όπου δηλώνει τα τελευταία F πλαίσια για τα οποία θα υπολογιστεί ο μέσος όρος τους. Ως τιμές δοκιμάσαμε τις εξής: 25, 50, 75, 100

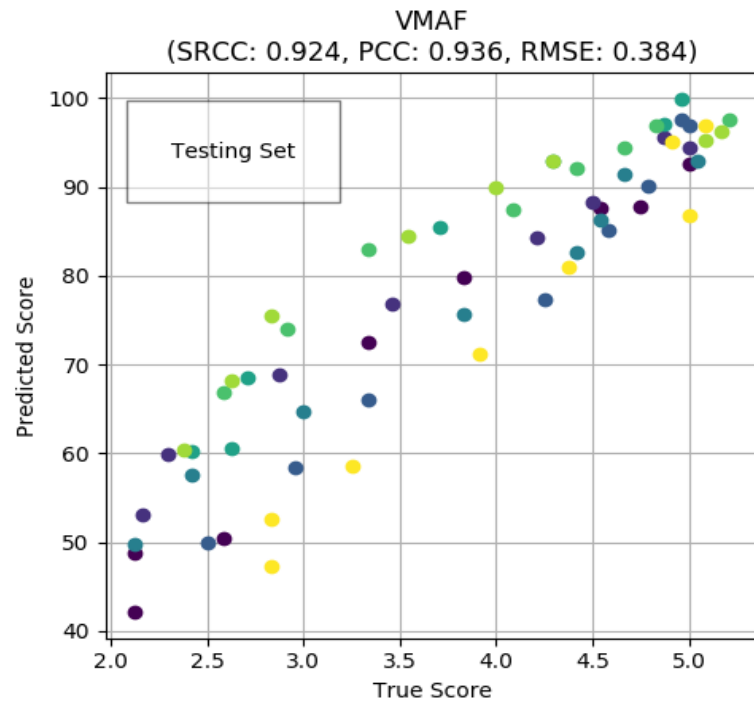
5.2.1. Αποτελέσματα με τη χρήση του NETFLIX Video Dataset

Στον Πίνακα 5.3 φαίνονται οι τιμές των SRCC, PCC και RMSE για τις διάφορες τιμές του F χρησιμοποιώντας ως μέθοδο χρονικής συγκέντρωσης την Μέσο όρο τελευταίων πλαισίων στο NETFLIX Video Dataset. Τα καλύτερα αποτελέσματα προκύπτουν όταν η μεταβλητή F πάρει την τιμή 50. Από την τιμή 25 έως την τιμή 50 παρατηρείται αύξηση των SRCC και PCC και μείωση του RMSE, αμέσως μετά όμως την τιμή 50 οι τιμές των SRCC και PCC μειώνονται ενώ ο RMSE αυξάνει.

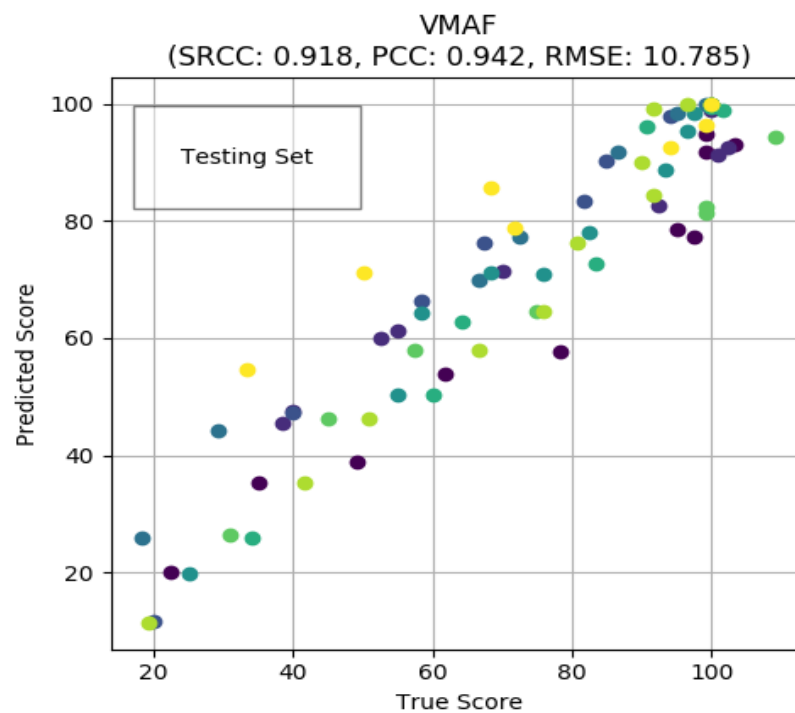
Πίνακας 5-3 Τιμές SRCC, PCC και RMSE για τις διάφορες τιμές της παραμέτρου F.

VMAF				
Μέθοδοι χρονικής συγκέντρωσης		SRCC	PCC	RMSE
Μέσος όρος τελευταίων πλαισίων	F = 25	0.907	0.926	11.937
	F = 50	0.918	0.942	10.785
	F = 75	0.918	0.941	10.86
	F = 100	0.915	0.934	11.482

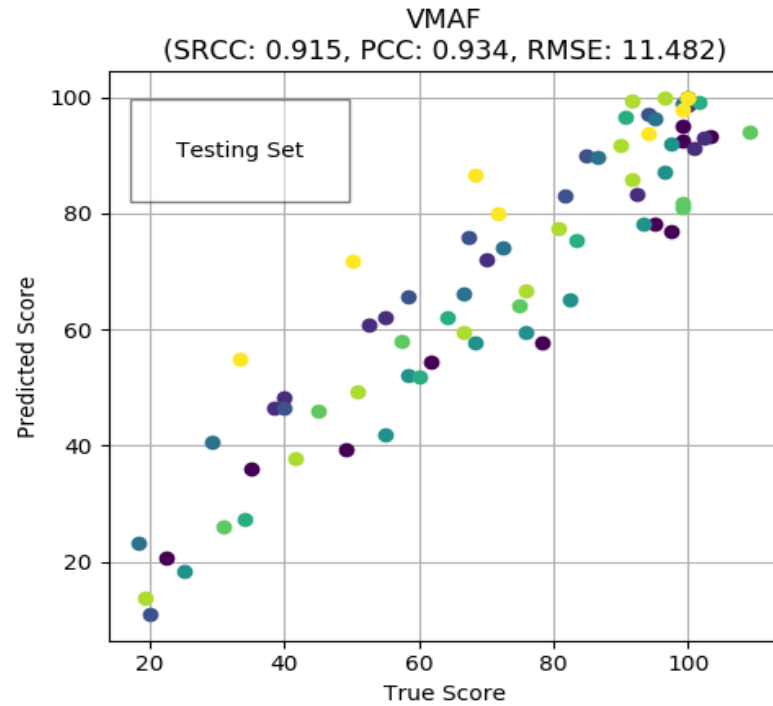
Οι Εικόνες 5.12, 5.13, 5.14 και 5.15 δείχνουν τα SRCC, PCC και RMSE αποτελέσματα της χρήσης του Μέσου όρου τελευταίων πλαισίων ως μέθοδο χρονικής συγκέντρωσης για τις διάφορες του F.



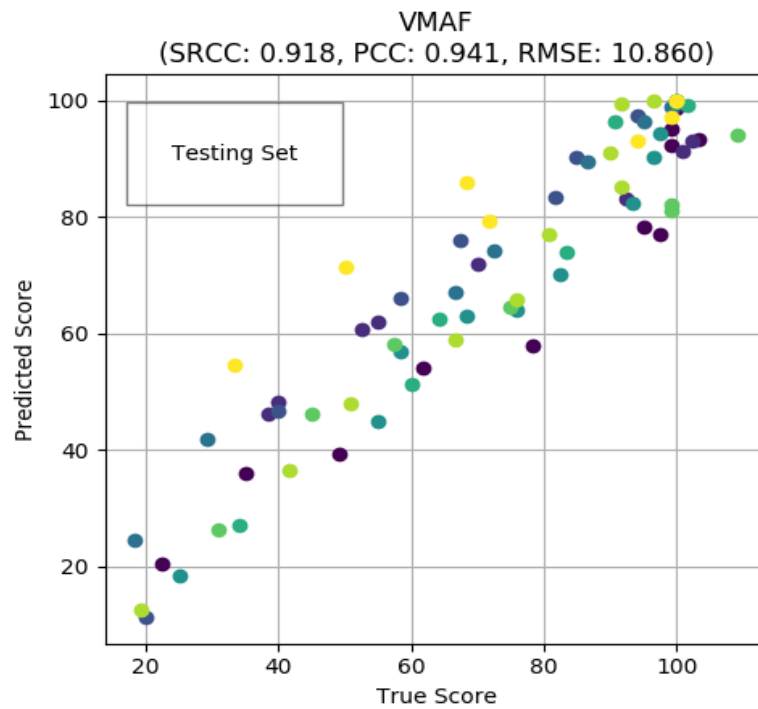
Εικόνα 5-12 Γράφημα αποτελεσμάτων VMAF scores με χρήση του Μέσου όρου τελευταίων πλαισίων με F=25



Εικόνα 5-13 Γράφημα αποτελεσμάτων VMAF scores με χρήση του Μέσου όρου τελευταίων πλαισίων με F=50



Εικόνα 5-14 Γράφημα αποτελεσμάτων VMAF scores με χρήση του Μέσου όρου τελευταίων πλαισίων με F=75



Εικόνα 5-15 Γράφημα αποτελεσμάτων VMAF scores με χρήση του Μέσου όρου τελευταίων πλαισίων με F=100

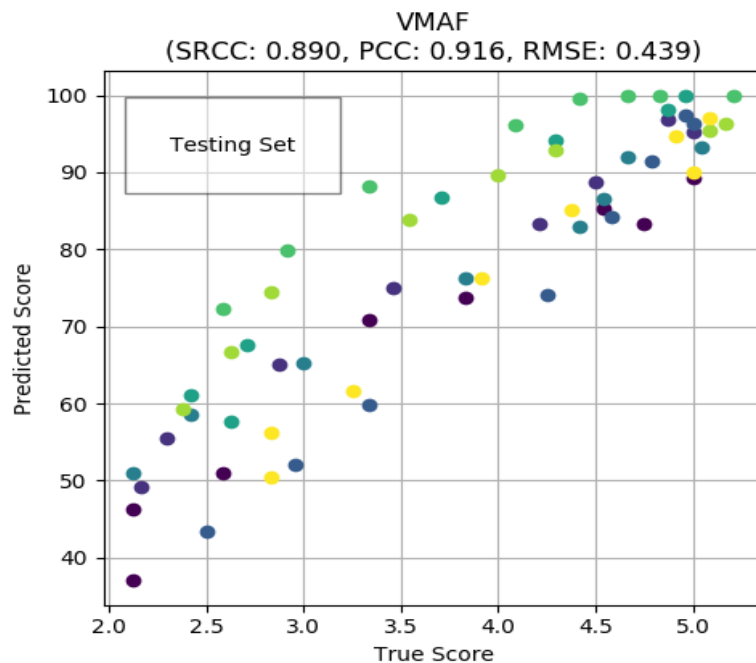
5.2.2. Αποτελέσματα με τη χρήση του VQEG HD3 Dataset

Στον Πίνακα 5.4 φαίνονται οι τιμές των SRCC, PCC και RMSE για τις διάφορες τιμές του F χρησιμοποιώντας ως μέθοδο χρονικής συγκέντρωσης τον Μέσο όρο τελευταίων πλαισίων στο VQEG HD3 Dataset. Τα καλύτερα αποτελέσματα προκύπτουν όταν η μεταβλητή F πάρει την τιμή 75.

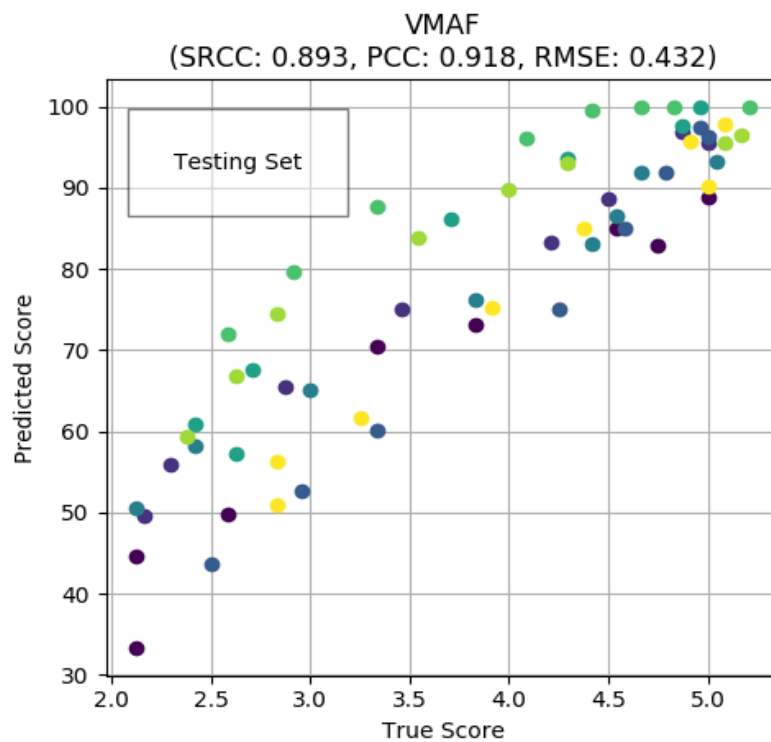
Πίνακας 5-4 Τιμές SRCC, PCC και RMSE για τις διάφορες τιμές της παραμέτρου F.

VMAF				
Μέθοδοι χρονικής συγκέντρωσης		SRCC	PCC	RMSE
Μέσος όρος τελευταίων πλαισίων	F = 25	0.890	0.916	0.439
	F = 50	0.893	0.918	0.432
	F = 75	0.909	0.929	0.401
	F = 100	0.913	0.932	0.392

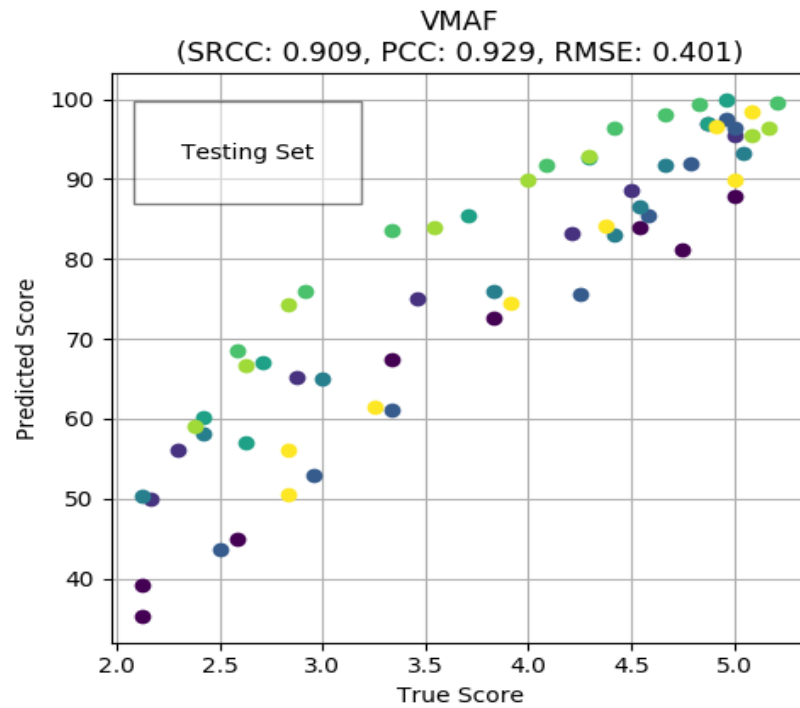
Οι Εικόνες 5.16, 5.17, 5.18 και 5.19 δείχνουν τα SRCC, PCC και RMSE αποτελέσματα της χρήσης του Μέσου όρου τελευταίων πλαισίων ως μέθοδο χρονικής συγκέντρωσης για τις διάφορες του F.



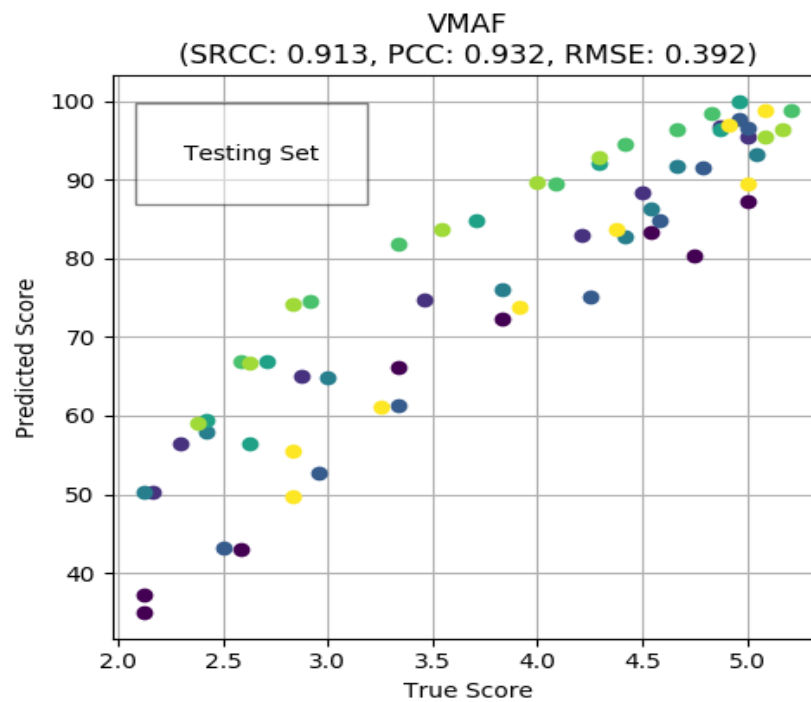
Εικόνα 5-16 Γράφημα αποτελεσμάτων VMAF scores με χρήση του Μέσου όρου τελευταίων πλαισίων με F=25



Εικόνα 5-17 Γράφημα αποτελεσμάτων VMAF scores με χρήση του Μέσου όρου τελευταίων πλαισίων με F=50



Εικόνα 5-18 Γράφημα αποτελεσμάτων VMAF scores με χρήση του Μέσου όρου τελευταίων πλαισίων με F=75



Εικόνα 5-19 Γράφημα αποτελεσμάτων VMAF scores με χρήση του Μέσου όρου τελευταίων πλαισίων με F=100

5.3. K-th percentile

Για τη χρήση της μεθόδου χρονικής συγκέντρωσης k-th percentile, δοκιμάσαμε διάφορες τιμές για την παράμετρο k όπου δηλώνει το ποσοστό k% των πλαισίων με τις χαμηλότερες τιμές. Ως k δοκιμάστηκαν οι τιμές 5, 10, 20, 25.

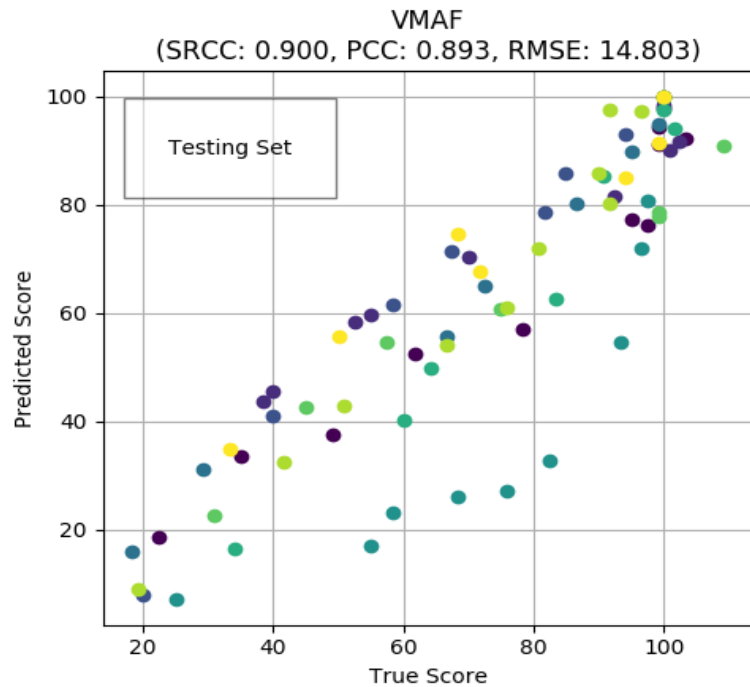
5.3.1. Αποτελέσματα με τη χρήση του NETFLIX Video Dataset

Στον Πίνακα 5.5 φαίνονται οι τιμές των SRCC, PCC και RMSE για τις διάφορες τιμές του k χρησιμοποιώντας ως μέθοδο χρονικής συγκέντρωσης την k-th percentile στο NETFLIX Video Dataset. Τα καλύτερα αποτελέσματα προκύπτουν όταν η μεταβλητή k πάρει την τιμή 25.

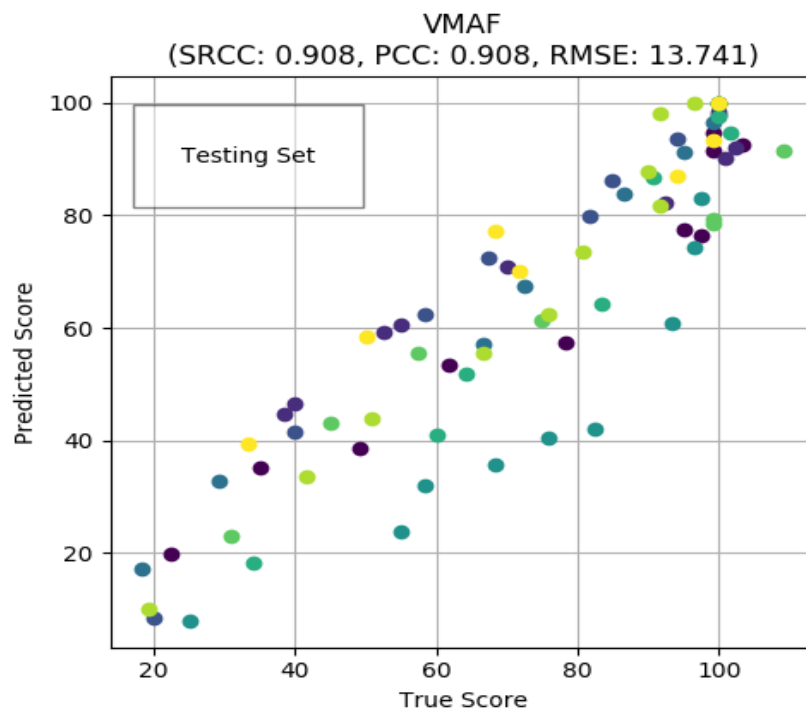
Πίνακας 5-5 Τιμές SRCC, PCC και RMSE για τις διάφορες τιμές της παραμέτρου k.

VMAF				
Μέθοδοι χρονικής συγκέντρωσης		SRCC	PCC	RMSE
k-th percentile	k = 5	0.900	0.893	14.803
	k = 10	0.908	0.908	13.741
	k = 20	0.907	0.911	13.422
	k = 25	0.918	0.934	11.528

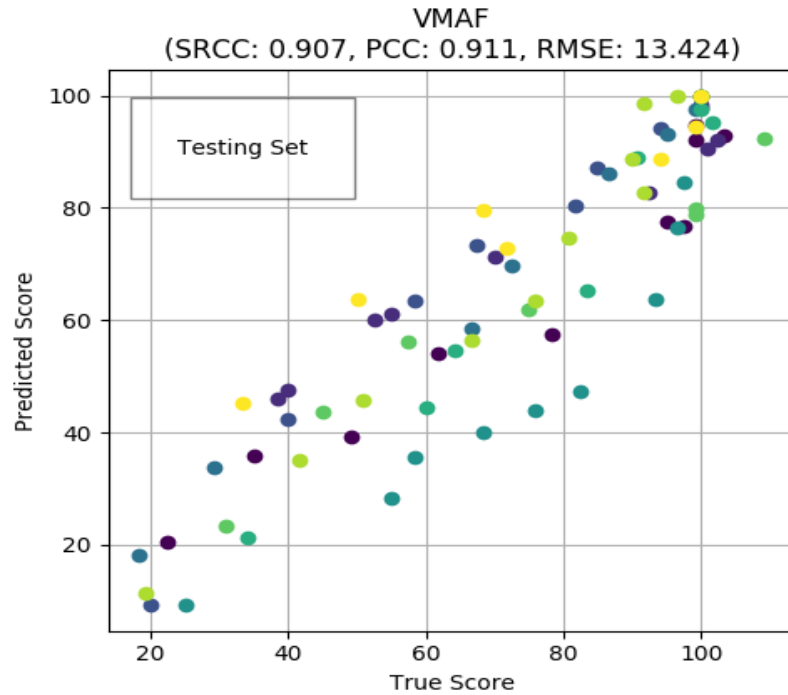
Οι Εικόνες 5.20, 5.21, 5.22 και 5.23 δείχνουν τα SRCC, PCC και RMSE αποτελέσματα της χρήσης της k-th percentile ως μέθοδο χρονικής συγκέντρωσης για τις διάφορες τιμές του k.



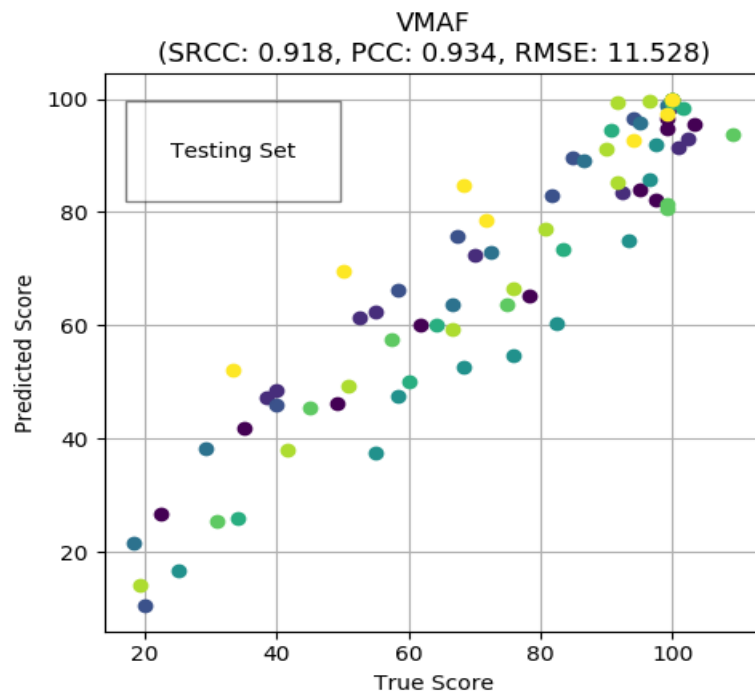
Εικόνα 5-20 Γράφημα αποτελεσμάτων VMAF scores με μέθοδο χρονικής συγκέντρωσης την k -th percentile με $k=5$



Εικόνα 5-21 Γράφημα αποτελεσμάτων VMAF scores με μέθοδο χρονικής συγκέντρωσης την k -th percentile με $k=10$



Εικόνα 5-22 Γράφημα αποτελεσμάτων VMAF scores με μέθοδο χρονικής συγκέντρωσης την k -th percentile με $k=20$



Εικόνα 5-23 Γράφημα αποτελεσμάτων VMAF scores με μέθοδο χρονικής συγκέντρωσης την k -th percentile με $k=25$

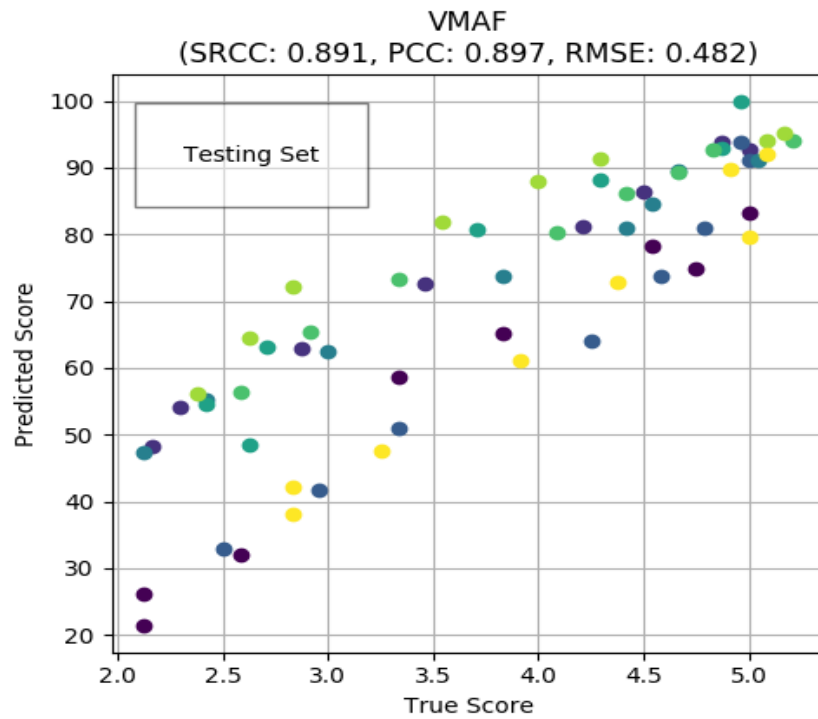
5.3.2. Αποτελέσματα με τη χρήση του VQEG HD3 Dataset

Στον Πίνακα 5.6 φαίνονται οι τιμές των SRCC, PCC και RMSE για τις διάφορες τιμές του k χρησιμοποιώντας ως μέθοδο χρονικής συγκέντρωσης την k -th percentile στο VQEG HD3 Dataset. Τα καλύτερα αποτελέσματα προκύπτουν όταν η μεταβλητή k πάρει την τιμή 25.

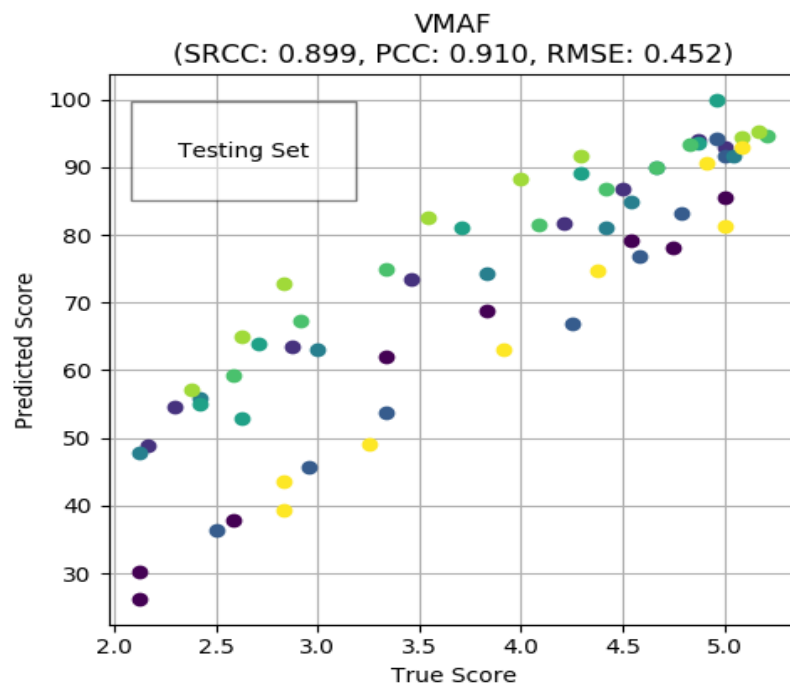
Πίνακας 5-6 Τιμές SRCC, PCC και RMSE για τις διάφορες τιμές της παραμέτρου k .

VMAF				
Μέθοδοι χρονικής συγκέντρωσης		SRCC	PCC	RMSE
k-th percentile	k = 5	0.891	0.897	0.482
	k = 10	0.899	0.910	0.452
	k = 20	0.909	0.925	0.410
	k = 25	0.924	0.936	0.384

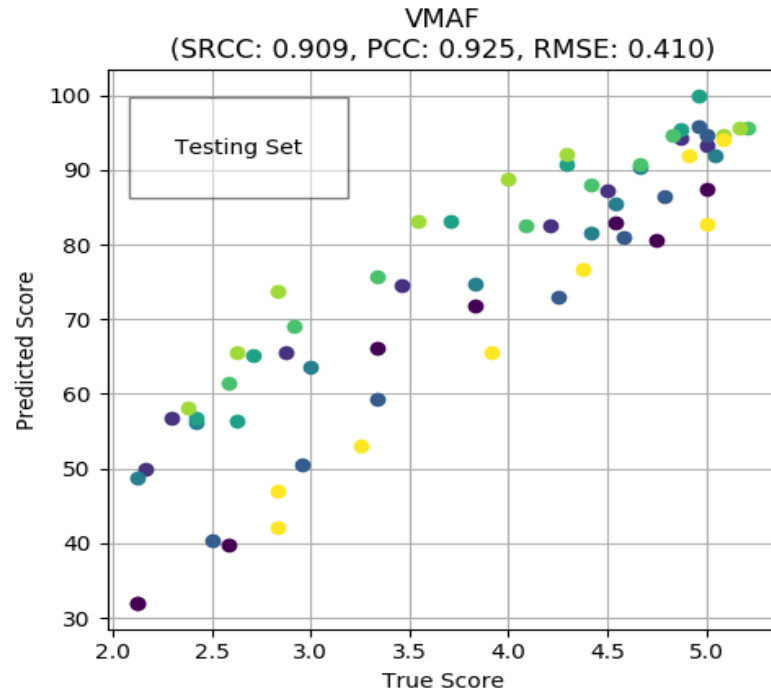
Οι Εικόνες 5.24, 5.25, 5.26 και 5.27 δείχνουν τα SRCC, PCC και RMSE αποτελέσματα της χρήσης της k -th percentile ως μέθοδο χρονικής συγκέντρωσης για τις διάφορες του k .



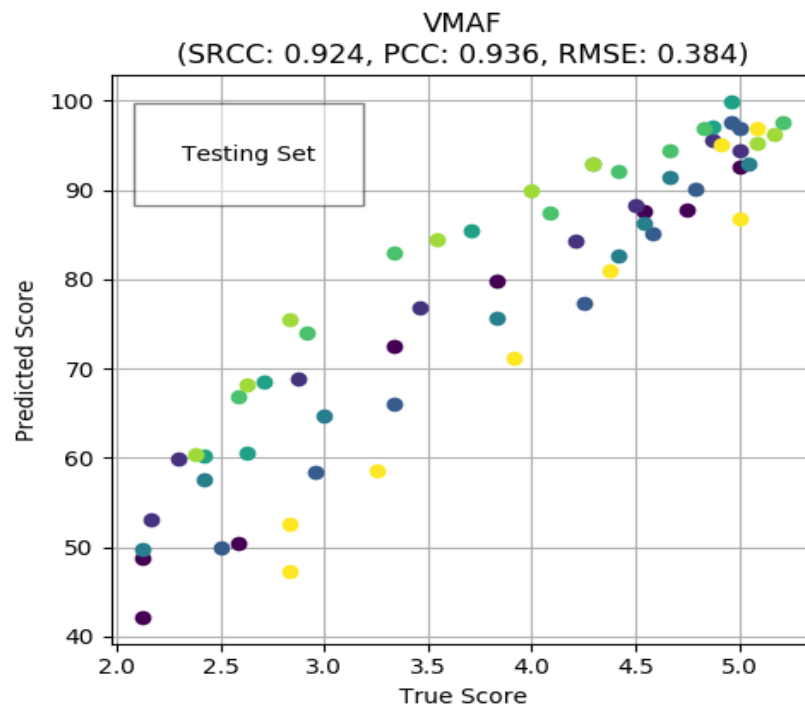
Εικόνα 5-24 Γράφημα αποτελεσμάτων VMAF scores με μέθοδο χρονικής συγκέντρωσης την k -th percentile με $k=5$



Εικόνα 5-25 Γράφημα αποτελεσμάτων VMAF scores με μέθοδο χρονικής συγκέντρωσης την k -th percentile με $k=10$



Εικόνα 5-26 Γράφημα αποτελεσμάτων VMAF scores με μέθοδο χρονικής συγκέντρωσης την k -th percentile με $k=20$



Εικόνα 5-27 Γράφημα αποτελεσμάτων VMAF scores με μέθοδο χρονικής συγκέντρωσης την k -th percentile με $k=25$

5.4. Μέγιστη, ελάχιστη τιμή και διάμεσος

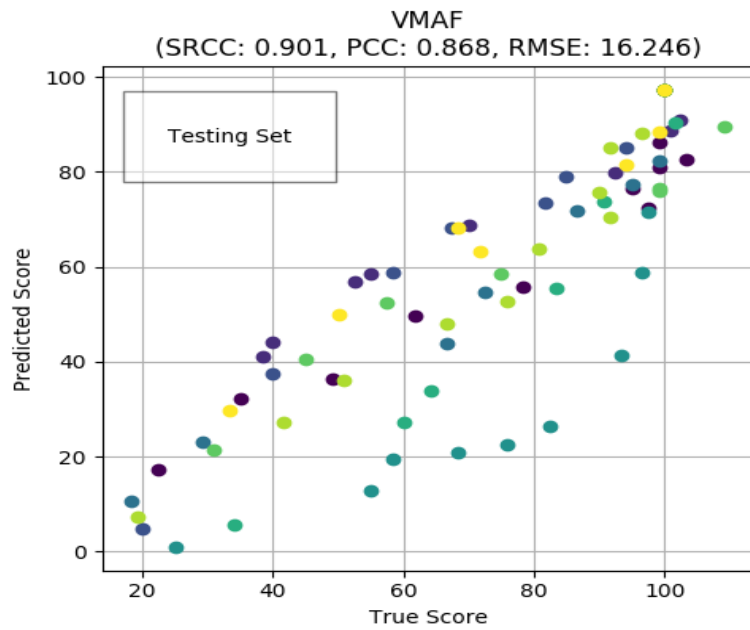
5.4.1. Αποτελέσματα με τη χρήση του NETFLIX Video Dataset

Στον Πίνακα 5.7 φαίνονται οι τιμές των SRCC, PCC και RMSE για το NETFLIX Video Dataset για τη χρήση των min, max και median ως μέθοδοι χρονικής συγκέντρωσης. Τα καλύτερα αποτελέσματα προκύπτουν για την median (διάμεσο) .

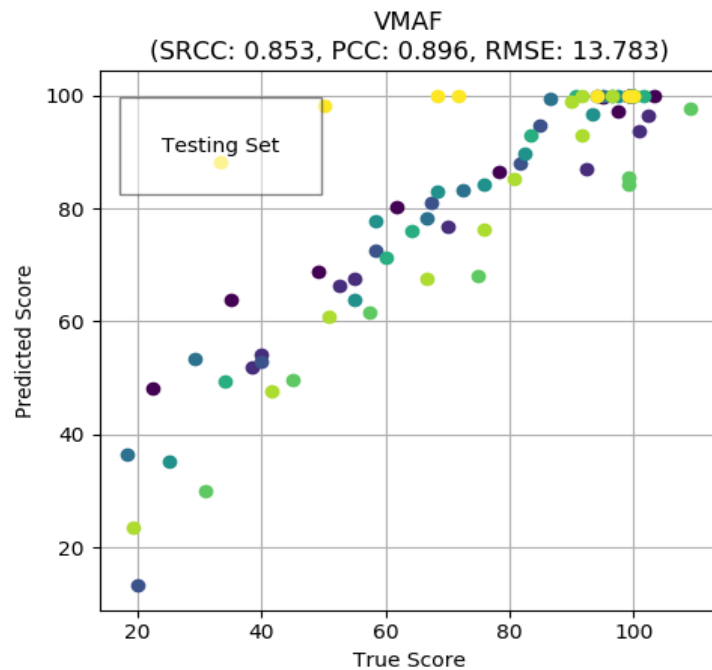
Πίνακας 5-7 Τιμές SRCC, PCC και RMSE για τις τιμές min, max και median.

VMAF			
Μέθοδοι χρονικής συγκέντρωσης	SRCC	PCC	RMSE
Min	0.901	0.868	16.246
Max	0.853	0.896	13.783
Median (διάμεσος)	0.912	0.929	11.974

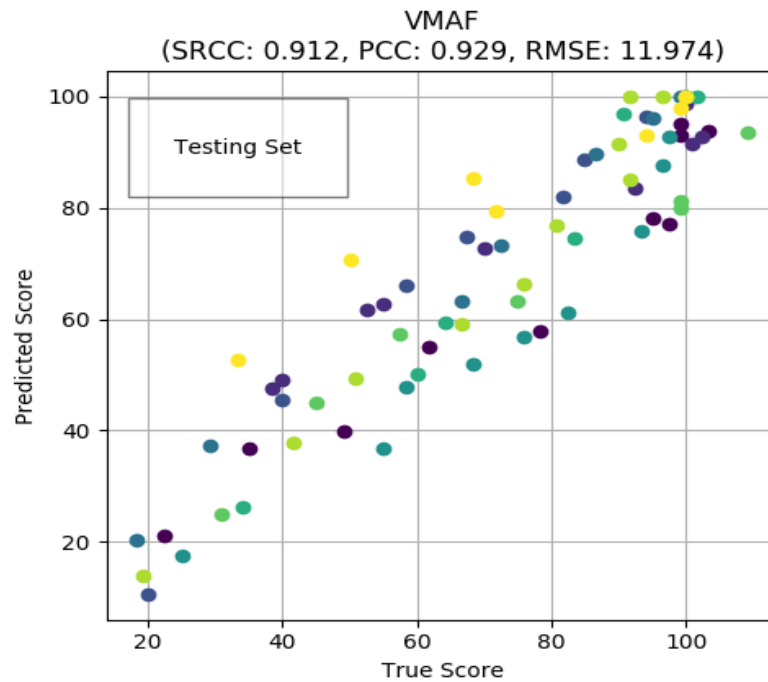
Οι Εικόνες 5.28, 5.29 και 5.30 δείχνουν τα SRCC, PCC και RMSE αποτελέσματα της χρήσης των ελάχιστη, μέγιστη τιμή και διάμεσος ως μέθοδοι χρονικής συγκέντρωσης αντίστοιχα.



Εικόνα 5-28 Γράφημα αποτελεσμάτων VMAF scores με μέθοδο χρονικής συγκέντρωσης την ελάχιστη τιμή.



Εικόνα 5-29 Γράφημα αποτελεσμάτων VMAF scores με μέθοδο χρονικής συγκέντρωσης τη μέγιστη τιμή.



Εικόνα 5-30 Γράφημα αποτελεσμάτων VMAF scores με μέθοδο χρονικής συγκέντρωσης την διάμεσο.

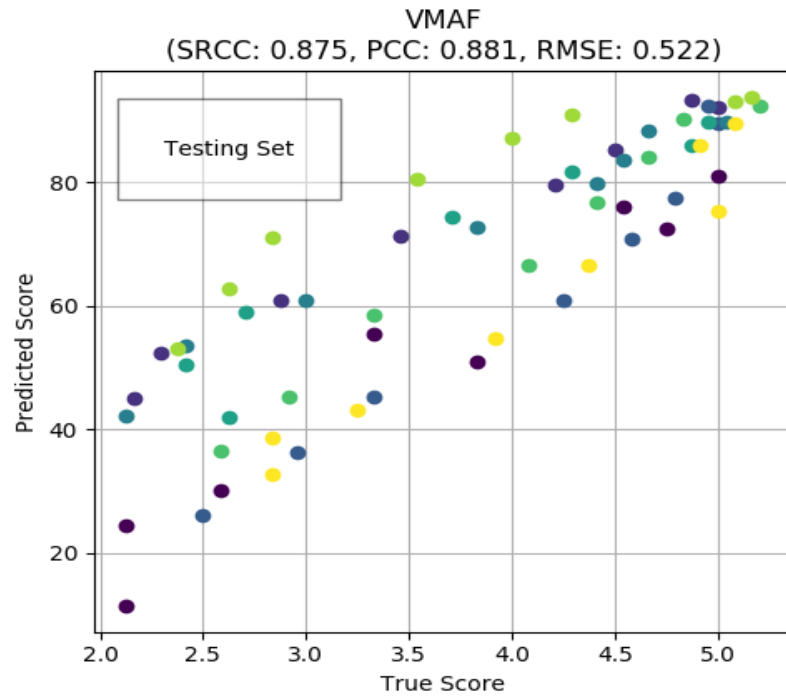
5.4.2. Αποτελέσματα με τη χρήση του VQEG HD3 Dataset

Στον Πίνακα 5.8 φαίνονται οι τιμές των SRCC, PCC και RMSE για το VQEG HD3 Dataset για τη χρήση των ελάχιστη, μέγιστη τιμή και διάμεσο ως μέθοδο χρονικής συγκέντρωσης. Τα καλύτερα αποτελέσματα προκύπτουν για την διάμεσο .

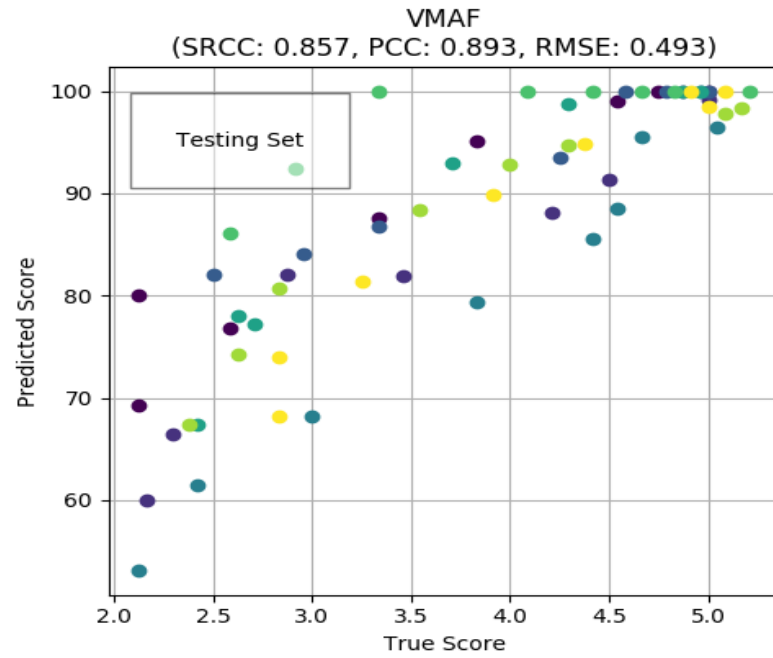
Πίνακας 5-8 Τιμές SRCC, PCC και RMSE για τις τιμές min, max και median.

VMAF			
Μέθοδοι χρονικής συγκέντρωσης	SRCC	PCC	RMSE
Min	0.875	0.881	0.522
Max	0.857	0.899	0.493
Median (διάμεσος)	0.930	0.938	0.378

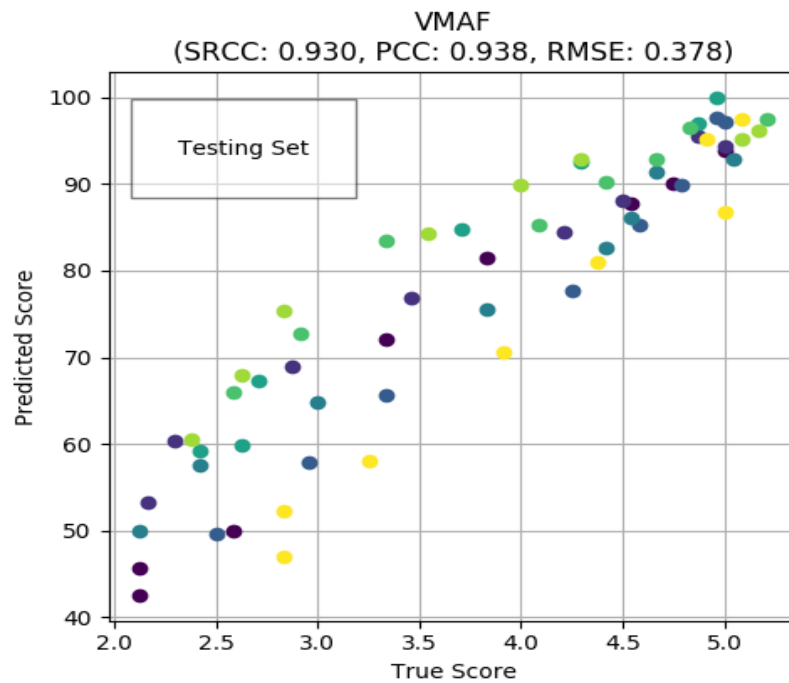
Οι Εικόνες 5.31, 5.32 και 5.33 δείχνουν τα SRCC, PCC και RMSE αποτελέσματα της χρήσης των Min , Max και Median ως μέθοδοι χρονικής συγκέντρωσης αντίστοιχα.



Εικόνα 5-31 Γράφημα αποτελεσμάτων VMAF scores με μέθοδο χρονικής συγκέντρωσης την ελάχιστη τιμή.



Εικόνα 5-32 Γράφημα αποτελεσμάτων VMAF scores με μέθοδο χρονικής συγκέντρωσης τη μέγιστη τιμή.



Εικόνα 5-33 Γράφημα αποτελεσμάτων VMAF scores με μέθοδο χρονικής συγκέντρωσης τη διάμεσο.

ΚΕΦΑΛΑΙΟ 6. ΣΥΜΠΕΡΑΣΜΑΤΑ

Η παρούσα εργασία εκπονήθηκε με σκοπό να αποτελέσει μία συνοπτική μελέτη του τρόπου εκτίμησης της αντιληπτής ποιότητας βίντεο από μετρικές επιπέδου πλαισίου/frame, όπως η μετρική VMAF. Στα πλαίσια αυτά, χρησιμοποιήθηκε η ανοιχτού κώδικα υλοποίηση της μεθόδου VMAF καθώς και αξιόπιστα αποθετήρια βίντεο, συνοδευόμενα από υποκειμενικές βαθμολογίες χρηστών. Εξετάστηκαν, σε ένα σύνολο πειραμάτων βασιζόμενα σε παλαιότερες δημοσιεύσεις [25] και [26], διαφορετικές μέθοδοι χρονικής συγκέντρωσης αποτελέσματος (pooling methods), για την εξαγωγή μίας ολικής τιμής αντιληπτής ποιότητας ενός βίντεο από το σύνολο των τιμών VMAF των πλαισίων του. Το βασικό συμπέρασμα των πειραμάτων είναι η επιβεβαίωση της υπόθεσης ότι η χρήση του απλού μέσου όρου είναι μια αισιόδοξη προσέγγιση, δηλαδή υποεκτιμά τη συνεισφορά των χαμηλών τιμών VMAF (πλαίσια με κακή ποιότητα) στην συνολική τιμή αντιληπτής ποιότητας βίντεο (μετρούμενη σε MOS). Σε αυτή την βάση, άλλες μέθοδοι συγκέντρωσης αποτελέσματος, όπως αυτές που προκύπτουν από το άθροισμα Minkowski, αποδείχθηκε ότι μπορούν να προσεγγίσουν καλύτερα τις υποκειμενικές μετρήσεις.

Από τα αποτελέσματα προέκυψε επίσης ότι η υποκειμενική βαθμολόγηση των χρηστών επηρεάζεται κατά κύριο λόγο από: το μέρος του βίντεο που είναι χαμηλής ποιότητας, καθώς και τα τελευταία δευτερόλεπτα προβολής του.

Μια συνολική ματιά στα πειράματα, δείχνει πως η μέθοδος χρονικής συγκέντρωσης αθροίσματος Minkowski κατάλληλα παραμετροποιημένη (όταν η παράμετρος Minkowski έχει την τιμή $p=8$) υπερέρχει όλων των άλλων μεθόδων χρονικής συγκέντρωσης.

Ένα επόμενο βήμα έρευνας είναι να διεξαχθεί μια βαθύτερη έρευνα σχετικά με την πιθανή σχέση μεταξύ του περιεχομένου του βίντεο που εξετάζεται την εκάστοτε φορά και της μεθόδου χρονικής συγκέντρωσης που χρησιμοποιείται με τη VMAF.

Είναι πιθανό διαφορετικές μέθοδοι χρονικής συγκέντρωσης να αποδίδουν καλύτερα σε συγκεκριμένου περιεχομένου βίντεο.

ΑΝΑΦΟΡΕΣ

- [1] F. Carvalho, T. Magedanz. "Telecommunication Systems and Technologies: Quality of Service in Telecommunication Networks", 2009.
- [2] Staelens, Nicolas, et al. "Measuring video quality in the network: from quality of service to user experience." 9th International Workshop on Video Processing and Consumer Electronics (VPQM 2015), 2015.
- [3] Girod B. "What's wrong with mean-squared error?" In A. B. Watson (Ed.), "Digital images and human vision" (pp. 207–220). Cambridge: MIT Press, 1993.
- [4] P. Teo, D. Heeger. "Perceptual image distortion. In Proceedings of SPIE human vision, visual processing and digital display V" (Vol. 2179, pp. 127–141), San Jose, February 1994.
- [5] Z. Wang, A. C. Bovik. "A universal image quality index. IEEE Signal Processing Letters", 9 (3), 81–84, 2002.
- [6] Z. Wang, A. C. Bovik. "Why is image quality assessment so difficult? In Proceedings of IEEE international conference on acoustics, speech, and signal processing" (ICASSP) (Vol. 4, pp. 3313–3316), Orlando, May 2002.
- [7] O. Nemethova, M. Ries, E. Siffel, M. Rupp. "Quality assessment for H.264 coded low-rate and low-resolution video sequences". In Third IASTED international conference on communications, Internet, and information technology (CIIT 2004) (pp. 136–140), St. Thomas, US Virgin Islands, 22–24 January 2004.
- [8] I. Richardson. H.264 and MPEG-4 video compression: "video coding for next-generation multimedia". New York: Wiley, 2004.

- [9] S. Wolf, M. Pinson, A. Webster, G. Cermak, E. P. Tweedy. "Objective and subjective measures of MPEG video quality". ITS/NTIA, Tech. Rep. ANSI T1A1.5/96-121, October 1996.
- [10] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Trans. on Image Proc.*, vol. 13, no. 4, pp. 600–612, 2004.
- [11] Z. Wang, L. Lu, and A. C. Bovik, "Video quality assessment based on structural distortion measurement," *Signal processing: Image communication*, vol. 19, no. 2, pp. 121–132, 2004.
- [12] M. H. Pinson, S. Wolf, "A new standardized method for objectively measuring video quality", *IEEE Transactions on Broadcasting*, vol. 50, no. 3, pp.312-322, September 2004.
- [13] NTIA, "VQM." http://www.its.bldrdoc.gov/n3/video/VQM_software.php (2008).
- [14] Z. Li, A. Aaron, I. Katsavounidis, A. Moorthy, M. Manohara, "Toward A Practical Perceptual Video Quality Metric", June 2016
- [15] Z. Li, C. Bampis, J. Novak, A. Aaron, K. Swanson, A. Moorthy, J. Cock "VMAF: The Journey Continues", October 2018.
- [16] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.
- [17] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Trans. Intelligent Systems and Technology*, vol. 2, no. 3, May 2011.
- [18] S. Li, F. Zhang, L. Ma, and K. N. Ngan, "Image quality assessment by separately evaluating detail losses and additive impairments," *IEEE Trans. Multimedia*, vol. 13, no. 5, pp. 935–949, 2011.
- [19] H. R. Sheikh and A. C. Bovik, "Image information and visual quality," *IEEE Trans. Image Process.*, vol. 15, no. 2, pp. 430–444, 2006.
- [20] M. Pinson and S. Wolf, "A new standardized method for objectively measuring video quality, *IEEE Trans. On Broad.*, vol. 50, no. 3, pp. 312-322, Sept. 2004.

- [21] K. Seshadrinathan and A. C. Bovik, "Motion tuned spatio-temporal quality assessment of natural videos," *IEEE Trans. Image Process.*, vol. 19, no. 2, pp. 335–350, Feb. 2010.
- [22] Z. Li, A. Aaron, I. Katsavounidis, A. Moorthy, and M. Manohara, "Toward a practical perceptual video quality metric [Online]. Available: <http://techblog.netflix.com/2016/06/towardpractical-perceptual-video.html>
- [23] HDTV Phase I Final Report [Online]. Available: <HTTPS://WWW.ITS.BLDRDOC.GOV/VOEG/PROJECTS/HDTV/HDTV.ASPX>.
- [24] H. R. Sheikh, M. F. Sabir, and A. C. Bovik, "A statistical evaluation of recent full reference image quality assessment algorithms," *IEEE Trans. Image Process.*, vol. 15, no. 11, pp. 3440–3451, Nov. 2006.
- [25] S. Rimac-Drlje, M. Vranjes, and D. Zagar, "Influence of temporal pooling method on the objective video quality evaluation," *IEEE International Symposium on Broadband Multimedia Systems and Broadcasting*, 2009.
- [26] M. Seufert, M. Slanina, S. Egger, M. Kottkamp, "To pool or not to pool: A comparison of temporal pooling methods for HTTP adaptive video streaming", *Fifth International Workshop on Quality of Multimedia Experience (QoMEX)*, Klagenfurt, Austria, July 2013.

ΠΑΡΑΡΤΗΜΑ

Η Netflix, στα πλαίσια του αλγορίθμου VMAF, προσφέρει στο Github τον κώδικα του αλγορίθμου και πολλές άλλες χρήσεις του όπως το να τρέξει μια βασική εντολή για αποτελέσματα VMAF, να τρέξει ένα αρχείο με πολλές εντολές για VMAF, ακόμα και training και testing μοντέλων τύπου VMAF ή/και διαφορετικών datasets από αυτά που προσφέρονται.

Εγκατάσταση προαπαιτούμενων σε Linux

Αρχικά απαραίτητο είναι να εγκατασταθούν ορισμένα από τα πακέτα που προσφέρει η Python, για να μπορέσει να τρέξει το πρόγραμμα. Η εντολή είναι:

```
sudo apt-get update -qq && sudo apt-get install -y pkg-config gfortran \ libhdf5-dev libfreetype6-dev liblapack-dev python python-setuptools \ python-dev python-pip python-tk
```

Στη συνέχεια αναβαθμίζουμε το pip και εγκαθιστούμε με αυτό τα πακέτα:

```
sudo -H pip install --upgrade pip
```

και

```
pip install --user numpy scipy matplotlib notebook pandas sympy nose \ scikit-learn scikit-image h5py
```

Και τέλος όσον αφορά τα Python πακέτα, τρέχουμε την παρακάτω εντολή και επανεκκινούμε το shell:

```
echo export PATH="$PATH:$HOME/.local/bin" >> ~/.bashrc
```

Εγκατάσταση των προγραμμάτων της VMAF

Αρχικά δημιουργούμε έναν νέο φάκελο που θα τον χρησιμοποιήσουμε για να κατεβάσουμε τα προγράμματα. Μόλις δημιουργηθεί, μπαίνουμε στον φάκελο αυτό και ανοίγουμε ένα νέο terminal και εκτελούμε τις παρακάτω εντολές για να κατεβάσουμε όλα τα αρχεία από το Github της Netflix:

git clone [HTTPS://GITHUB.COM/NETFLIX/VMAF](https://github.com/netflix/vmaf)

και στη συνέχεια μπαίνουμε με cd στον φάκελο vmaf και εκτελούμε τις εντολές:

```
git submodule update --init --recursive
```

και

```
make
```

Με την παρακάτω βασική εντολή για να προσθέσουμε τα paths 2 φακέλων στην μεταβλητή PYTHONPATH:

Για Linux (Ubuntu):

```
echo          export\PYTHONPATH="$(pwd)/python/src:$(pwd)/surreal/python/src:\
PYTHONPATH" >> ~/.bashrc
```

και

```
source ~/.bashrc
```

Για MacOS εκτελούμε τις ίδιες εντολές αντικαθιστώντας όπου bashrc με bash_profile.

Βασική χρήση των δυνατοτήτων του πακέτου VMAF

Αφού ολοκληρώσουμε την εγκατάσταση των παραπάνω, προκειμένου να επαληθεύσουμε αν έγιναν σωστά όλα τα βήματα, τρέχουμε στο τερματικό την παρακάτω εντολή:

```
./unittest
```

Αν δεν μας επιστρέψει κανένα error, τότε μπορούμε να συνεχίσουμε κανονικά.

Για να τρέξουμε το VMAF μεταξύ του original video και του παραμορφωμένου video τρέχουμε την παρακάτω εντολή:

```
./run_vmaf format width height reference_path disorted_path [--out-fmt output_format]
```

ακολουθούμενο από το format που θα είναι γυν420p, γυν422p, γυν444p, γυν420p10le, γυν422p10le ή γυν444p10le, το width και το height που θα είναι το πλάτος και το ύψος του video σε pixel, τα ref_path και dis_path που θα είναι τα paths του original video και του διαστρεβλωμένου video και προαιρετικά επιλογή format εξόδου των αποτελεσμάτων προσθέτοντας --out-fmt και μετά text, json ή xml, επιλογή κάποιου από της έτοιμες μεθόδους χρονικής συγκέντρωσης για τον τρόπο υπολογισμού της τιμής του vmaf προσθέτοντας --pool και μετά για παράδειγμα mean, καθώς και την επιλογή εμφάνισης των αποτελεσμάτων σε πιο γραφική μορφή προσθέτοντας --local-explain.

Άλλος ένας τρόπος για να τρέξουμε το VMAF είναι μέσω ενός αρχείου που κάθε γραμμή του θα περιέχει το παρακάτω format:

```
format width height ref_path dis_path
```

όπου καθεμιά από τις παραμέτρους είναι ίδια με προηγουμένως και στη συνέχεια τρέχουμε το πρόγραμμα ως εξής:

```
./run_vmaf_in_batch input_file [--out-fmt out_format] [--parallelize]
```

όπου το input_file θα είναι το path του αρχείου που δημιουργήσαμε σύμφωνα με το παραπάνω format.

Επίσης το πακέτο προσφέρει και επιπλέον επιλογές στις παραπάνω 2 χρήσεις. Μια από αυτές είναι ο υπολογισμός της τιμής του VMAF σε περίπτωση που έχουμε οθόνη κινητού όπου τότε προσθέτουμε στις παραπάνω 2 εντολές το --phone-model. Άλλη μια επιλογή είναι η χρήση και άλλου μοντέλου όπως το μοντέλο υπολογισμού VMAF σε 4K οθόνη προσθέτοντας --model και στη συνέχεια το path που βρίσκεται το μοντέλο. Τέλος άλλη μια επιλογή είναι η χρήση διαστήματος

εμπιστοσύνης των τιμών προσθέτοντας --ci στις εντολές.

Προχωρημένη χρήση του πακέτου VMAF

Η VMAF ακολουθεί μια προσέγγιση βασισμένη στην μηχανική μάθηση για να εξαγάγει πρώτα ένα πλήθος χαρακτηριστικών που σχετίζονται με την ποιότητα (ή στοιχειωδών μετρήσεων) από ένα παραμορφωμένο βίντεο και το βίντεο αναφοράς του, ακολουθούμενο από τη ανάμιξή τους σε μια τελική βαθμολογία ποιότητας χρησιμοποιώντας έναν μη γραμμικό regressor (π.χ. έναν SVM), εξ ου και το όνομα "Video Multi-method Assessment Fusion". . Για παράδειγμα, το μοντέλο καταλόγου περιέχει έναν αριθμό προ-εκπαιδευμένων μοντέλων, τα οποία μπορούν να φορτωθούν από τις προαναφερθείσες εντολές.

Εκτός από τις βασικές εντολές, το πακέτο VMAF παρέχει επίσης ένα πλαίσιο που επιτρέπει σε κάθε χρήστη να εκπαιδεύσει το δικό του μοντέλο αξιολόγησης της αντιληπτικής ποιότητας. Για να επιτευχθεί αυτό τα βήματα είναι τα εξής:

-Δημιουργία ενός dataset σύμφωνα με το format του αρχείο example_dataset.py που βρίσκεται στον φάκελο resource/example.

-Επικύρωση του dataset χρησιμοποιώντας είτε το ίδιο το VMAF είτε κάποια άλλη μετρική (PSNR, SSIM ή MS-SSIM), εκτελώντας την εντολή:

```
./run_testing quality_type dataset_file [--vmaf-model model_path] [--cache-result] [--parallelize]
```

όπου quality_type θα είναι VMAF, PSNR, SSIM ή MS-SSIM, dataset_file θα είναι το πλήρες path του αρχείου που δημιουργήσαμε στο προηγούμενο βήμα.

Εκπαίδευση ενός νέου μοντέλου εκτελώντας την εντολή:

```
./run_vmaf_training dataset_file feature_file model_param_file out_model_file [--cache-result] [--parallelize]
```

όπου feature_file θα είναι το path του αρχείου που θα περιέχει τα χαρακτηριστικά του μοντέλου, model_param_file θα είναι το path του αρχείου που θα περιέχει τον τύπο και τις παραμέτρους του μοντέλου και out_model_file θα είναι το αρχείο του μοντέλου μετά την εκπαίδευση το οποίο μπορεί μετά να χρησιμοποιηθεί με τις εντολές της βασικής χρήσης.

Επιπρόσθετα άλλες δύο λειτουργίες που παρέχει το πακέτο είναι:

-Χρήση τυχαίων υποκειμενικών μοντέλων για την εκπαίδευση και το τεστάρισμα ενός μοντέλου γράφοντας στις εντολές ./run_vmaf_training και ./run_testing --subj-model και μετά το υποκειμενικό μοντέλο π.χ. DMOS, MLE, κ.α.

-Διασταυρωμένη επικύρωση παραμέτρων και μοντέλων όπου δημιουργεί ένα μοντέλο βάση ενός αρχείου παραμέτρων και το χρησιμοποιεί σε ένα νέο dataset και εξάγει τιμές συσχέτισης.

Υλοποίηση των πειραμάτων

Για την υλοποίηση των πειραμάτων χρειάστηκε να τροποποιήσουμε το αρχείο stats.py έτσι ώστε να υλοποιήσουμε κάποιες από τις μεθόδους χρονικής συγκέντρωσης που δεν υπήρχαν, και στη συνέχεια το αρχείο run_testing.py με σκοπό να τις προσθέσουμε ονομαστικά στη λίστα από την οποία διαβάζει τις υπάρχουσες μεθόδους χρονικής συγκέντρωσης.

Μετά όσον αφορά τα εκάστοτε datasets, βάζαμε τα βίντεο στον φάκελο /path/ και η τελικά για να πάρουμε αποτελέσματα τρέχουμε την εντολή

```
./run_testing quality_type dataset_file --cache-result --parallelize --pool [pooling_method]
```

Αλλάζοντας κάθε φορά το dataset_file και την μέθοδο χρονικής συγκέντρωσης.

Ο λόγος που δεν κάνουμε εκ νέου training με το εκάστοτε καινούριο dataset είναι διότι επικεντρωνόμαστε στην προσωρινή συγκέντρωση των VMAF τιμών. Βασιζόμαστε επομένως, στην εκπαίδευση που έχει ήδη γίνει με τον τρόπο που αναφέρθηκε στο κεφάλαιο 3, και δοκιμάζουμε διαφορετικές μεθόδους χρονικής συγκέντρωσης, υποθέτοντας ότι το μοντέλο που έχουμε στην διάθεσή μας είναι το κατάλληλο.

