

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΑΤΡΩΝ - ΠΟΛΥΤΕΧΝΙΚΗ ΣΧΟΛΗ
ΤΜΗΜΑ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΤΕΧΝΟΛΟΓΙΑΣ ΥΠΟΛΟΓΙΣΤΩΝ



ΠΑΝΕΠΙΣΤΗΜΙΟ
ΠΑΤΡΩΝ
UNIVERSITY OF PATRAS

ΤΟΜΕΑΣ: Η/Υ

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

του φοιτητή του Τμήματος Ηλεκτρολόγων Μηχανικών και Τεχνολογίας
Υπολογιστών της Πολυτεχνικής Σχολής του Πανεπιστημίου Πατρών

ΑΓΓΕΛΟΥ ΚΑΡΔΟΥΤΣΟΥ ΤΟΥ ΑΠΟΣΤΟΛΟΥ

ΑΡΙΘΜΟΣ ΜΗΤΡΩΟΥ: 1059372

Θέμα

Ανάπτυξη συστήματος μεικτής πραγματικότητας για υποστήριξη
ατόμων με δυσκολίες όρασης

Επιβλέπων

Νικόλαος Αβούρης

Αριθμός Διπλωματικής Εργασίας: XXXX

Πάτρα, Οκτώβριος 2023

ΠΙΣΤΟΠΟΙΗΣΗ

Πιστοποιείται ότι η διπλωματική εργασία με θέμα

Ανάπτυξη συστήματος μεικτής πραγματικότητας για υποστήριξη
ατόμων με δυσκολίες όρασης

του φοιτητή του Τμήματος Ηλεκτρολόγων Μηχανικών και
Τεχνολογίας Υπολογιστών

Άγγελου Καρδούτσου του Απόστο

(Α.Μ.: 1059372)

παρουσιάστηκε δημόσια και εξετάστηκε στο τμήμα Ηλεκτρολόγων
Μηχανικών και Τεχνολογίας Υπολογιστών στις

XX/10/2023

Ο Επιβλέπων

Ο Διευθυντής του Τομέα

Νικόλαος Αβούρης
Καθηγητής

Ονοματεπώνυμο Διευθυντή Τομέα
Βαθμίδα Διευθυντή Τομέα

Στοιχεία διπλωματικής εργασίας

Θέμα: Ανάπτυξη συστήματος μεικτής πραγματικότητας για
υποστήριξη ατόμων με δυσκολίες όρασης

Φοιτητής: Άγγελος Καρδούτσος του Απόστο

Ομάδα επίβλεψης

Νικόλαος Αβούρης
Βαθμίδα και Ονοματεπώνυμο Συνεπιβλέποντα
Ονοματεπώνυμο Διδακτορικού Φοιτητή

Περίοδος εκπόνησης της εργασίας:
Μάιος 2023 - Οκτώβριος 2023

Η εργασία αυτή γράφτηκε στο \LaTeX και χρησιμοποιήθηκε η
γραμματοσειρά GFS Didot του Greek Font Society.

Περίληψη

Η εργασία αυτή ασχολείται με ένα ιδιαίτερα ενδιαφέρον ζήτημα στο χώρο της επεξεργασίας σημάτων και εικόνων, την ανάλυση (resolution). Παρόλο που σήμερα στον κόσμο μας έχουμε καταφέρει να δημιουργήσουμε συσκευές με μεγάλη ευαισθησία καταγραφής, μεγάλο χώρο αποθήκευσης καθώς και υψηλούς ρυθμούς μετάδοσης και επεξεργασίας δεδομένων, εντούτοις υπάρχουν εφαρμογές όπου η φύση τους είναι τέτοια που δε μας επιτρέπει να επωφεληθούμε σε μεγάλο βαθμό από την πρόοδο που έχει σημειωθεί. Μια τέτοια εφαρμογή είναι οι θερμικές εικόνες και θα δούμε στη συνέχεια της εργασίας τους λόγους εκείνους που την καθιστούν “ιδιαιτέρη”.

Ευχαριστίες

Όσο κι αν φαίνεται σαν ατομική δουλειά η παρούσα εργασία, στην πραγματικότητα βοήθησαν αρκετοί άνθρωποι (ο καθένας με το δικό του τρόπο) για να ολοκληρωθεί.

ΠΕΡΙΕΧΟΜΕΝΑ

Κατάλογος σχημάτων	xiii
Κατάλογος πινάκων	xv
1 Εισαγωγή	1
2 Θεωρητικό και Τεχνολογικό Υπόβαρο	3
2.1 Η υπέρυθη ακτινοβολία	3
3 Η υλοποίηση	5
3.1 Μοντελοποίηση προβλήματος	5
4 Αξιολόγηση	7
5 Αποτελέσματα	9
6 Μελλοντικές Εφαρμογές	11
Βιβλιογραφία	13

ΚΑΤΑΛΟΓΟΣ ΣΧΗΜΑΤΩΝ

2.1 Το φάσμα της Ηλεκτρομαγνητικής Ακτινοβολίας	4
---	---

ΚΑΤΑΛΟΓΟΣ ΠΙΝΑΚΩΝ

ΚΕΦΑΛΑΙΟ

1

ΕΙΣΑΓΩΓΗ

Γενική Εισαγωγή Δομή Εργασίας

Η εργασία αυτή έχει γίνει προσπάθεια να γραφεί σε ανεξάρτητα κεφάλαια, τα οποία θα δώσουν στον αναγνώστη τις απαιτούμενες γνώσεις ώστε να καταλάβει σε βάθος τις τεχνικές που χρησιμοποιούνται. Σε κάθε κεφάλαιο γίνεται αναλυτική παρουσίαση των τεχνικών καθώς και του υπόβαθρου που πρέπει να έχει κάποιος ώστε τις κατανοήσει, ωστόσο θεωρείται πως ο αναγνώστης έχει ήδη κάποιες γνώσεις στο χώρο της επεξεργασίας σήματος και εικόνας. Έτσι, βασικές έννοιες και μηχανισμοί της ανωτέρω περιοχής θα θεωρούνται δεδομένοι και δε θα γίνει κάποια ανάλυσή τους στο κείμενο αυτό, εκτός αν κρίνεται απαραίτητο.

ΚΕΦΑΛΑΙΟ

2

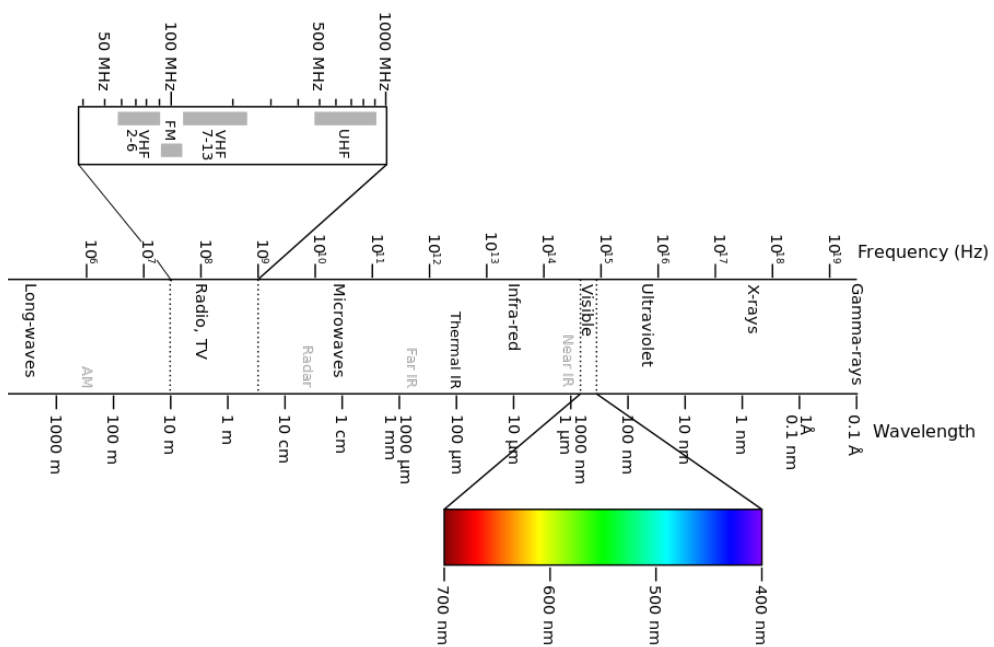
ΘΕΩΡΗΤΙΚΟ ΚΑΙ ΤΕΧΝΟΛΟΓΙΚΟ ΥΠΟΒΑΡΟ

ΘΕΩΡΙΑ

Blindness - Temporary solutions - Accessibility Extended reality Hololens -
Description of the device Hololens features Unity

2.1 Η υπέρυθρη ακτινοβολία

Όπως γνωρίζουμε από την επιστήμη της Φυσικής, όταν ένα σώμα ζεσταθεί αρκετά (π.χ. $> 100\text{ C}$) αρχίζει και εκπέμπει ακτινοβολία και στο οπτικό φάσμα (πέραν του υπέρυθρου), ενώ αποκτά ταυτόχρονα μια ερυθρή όψη. Κατά το φαινόμενο αυτό, η ακτινοβολία που εκπέμπεται ανήκει τόσο στο υπέρυθρο, όσο και στο ορατό φάσμα της ηλεκτρομαγνητικής ακτινοβολίας και μάλιστα το χρώμα που αποκτά το σώμα συνδέεται άμεσα με τη θερμοκρασία στην οποία βρίσκεται. Οι περιοχές της υπέρυθρης και ορατής ακτινοβολίας, καταλαμβάνουν τις ζώνες $1mm - 700nm$ και $700nm - 390nm$ αντίστοιχα, όπως φαίνεται και στο σχήμα 2.1.



Σχήμα 2.1: Το φάσμα της Ηλεκτρομαγνητικής Ακτινοβολίας

ΚΕΦΑΛΑΙΟ

3

Η ΥΛΟΠΟΙΗΣΗ

ΥΛΟΠΟΙΗΣΗ

3.1 Μοντελοποίηση προβλήματος

ΠΡΙΝ δούμε στην πράξη πώς λειτουργούν οι τεχνικές super resolution θα πρέπει να περιγράψουμε το πρόβλημα με το οποίο θα δουλέψουμε με μαθηματικούς όρους. Η μοντελοποίηση αυτή θα μας επιτρέψει να χρησιμοποιήσουμε μαθηματικά εργαλεία και τεχνικές και να ορίσουμε σε μια αυστηρή “γλώσσα” (αυτή των μαθηματικών) τις λειτουργίες που επιτελούνται από κάθε μέθοδο, ούτως ώστε να πετύχουμε το τελικό αποτέλεσμα.

Ξεκινώντας, θα πρέπει να περιγράψουμε τον τρόπο με τον οποίο λαμβάνουμε εικόνες χαμηλής ανάλυσης από μια φυσική σκηνή μέσω μιας κάμερας. Η κάμερα σαν όργανο καταγραφής εισάγει κάποιες ατέλειες, όπως είδαμε στο προηγούμενο κεφάλαιο. Τέτοιες ατέλειες μπορεί να είναι σφάλματα καταγραφής από τον αισθητήρα της κάμερας, θόλωμα λόγω αστοχιών των οπτικών στοιχείων κλπ. Επιπλέον, αν λάβουμε διαδοχικές εικόνες μιας φυσικής σκηνής, οι εικόνες αυτές περιμένουμε να έχουν κάποιες μετατοπίσεις ως προς αυτό που απεικονίζουν. Οι μετατοπίσεις αυτές μπορεί να οφείλονται είτε στον ανθρώπινο παράγοντα (που χειρίζεται την κάμερα) είτε στο αντικείμενο της φυσικής σκηνής που μπορεί να μην είναι σταθερό. Στην εικόνα που λαμβάνεται τελικά, λαμβάνονται δείγματα σε χαμηλή χωρική συχνότητα και λόγω ατελειών του οργάνου μπορεί να έχουμε και παρουσία θορύβου. Για μια εικόνα υψηλής ανάλυσης X (την οποία θα

προσπαθήσουμε να ανακατασκευάσουμε), μπορούμε να αναπτύξουμε το παραπάνω μοντέλο ως εξής:

$$\mathbf{Y}_i = \mathbf{S}_i \mathbf{T}_i \mathbf{H}_i \mathbf{X} + \mathbf{n}_i \quad (3.1)$$

όπου ορίζουμε για την i -οστή εικόνα χαμηλής ανάλυσης \mathbf{Y}_i τους τελεστές που ενεργούν στην \mathbf{X} :

- \mathbf{S}_i για υποδειγματοληψία,
- \mathbf{T}_i για μετατόπιση,
- \mathbf{H}_i για θόλωμα (blurring),
- \mathbf{n}_i για προσθετικό θόρυβο.

Οι υποθέσεις που κάνουμε για το παραπάνω πρόβλημα είναι ότι το blurring είναι ίδιο σε όλο το χώρο και είναι γνωστό στον αλγόριθμο super resolution, ο θόρυβος είναι λευκός Gaussian με την ίδια διασπορά σε όλες τις εικόνες χαμηλής ανάλυσης και ότι ο γεωμετρικός μετασχηματισμός αφορά μόνο την καθολική μετατόπιση.

Στην περίπτωση που θεωρήσουμε αμελητέο θόρυβο και θόλωμα, τα παραπάνω βήματα αρκούν για να υπολογίσουμε την εικόνα υψηλής ανάλυσης \mathbf{X} , την οποία λαμβάνουμε σαν αποτέλεσμα του αλγορίθμου.

Data: $\mathbf{dx}, \mathbf{dy}, \mathbf{Y}_i, N, W, \mathbf{H}, S$

Output: High resolution reconstructed \mathbf{X}

$\mathbf{X} \leftarrow 0;$

for $n \leftarrow 1$ to N do

$\tilde{\mathbf{Y}} \leftarrow \mathbf{Y}_i;$
 $i \leftarrow 1 : W/S;$
 $j \leftarrow 1 : H/S;$
 $\mathbf{px} = i * S + \mathbf{dx}_n;$
 $\mathbf{py} = j * S + \mathbf{dy}_n;$
 $\mathbf{X}_{\mathbf{px}, \mathbf{py}} = \tilde{\mathbf{Y}}_{i,j};$

end

Αλγόριθμος 1: Ανακατασκευή shift-add fusion

Συγκρίνοντας τη μέθοδο αυτή με την ανακατασκευή μέσω του μετασχηματισμού Fourier, παρατηρούμε την απλότητα και την αποδοτικότητα της καθώς βρίσκει την εικόνα υψηλής ανάλυσης μόνο με μετακινήσεις pixel.

ΚΕΦΑΛΑΙΟ

4

ΑΞΙΟΛΟΓΗΣΗ

ΑΞΙΟΛΟΓΗΣΗ

ΚΕΦΑΛΑΙΟ

5

ΑΠΟΤΕΛΕΣΜΑΤΑ

ΑΠΟΤΕΛΕΣΜΑΤΑ ΑΞΙΟΛΟΓΗΣΗΣ ΚΑΙ ΣΧΟ-
ΛΙΑ

ΚΕΦΑΛΑΙΟ

6

ΜΕΛΛΟΝΤΙΚΕΣ ΕΦΑΡΜΟΓΕΣ

ΒΕΛΤΙΩΣΕΙΣ ΚΑΙ ΠΡΟΕΚΤΑΣΕΙΣ

ΒΙΒΛΙΟΓΡΑΦΙΑ

- [1] S. C. Park, M. K. Park, and M. G. Kang, “Super-resolution image reconstruction: A technical overview,” *IEEE Signal Processing Magazine*, vol. 20, no. 3, pp. 21–36, May 2003.
- [2] D. Keren, S. Peleg, and R. Brada, “Image sequence enhancement using subpixel displacements,” in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, June 1988, pp. 742–746.
- [3] R. Hardie, K. Barnard, and E. Armstrong, “Joint MAP registration and high resolution image estimation using a sequence of undersampled images,” *IEEE Transactions on Image Processing*, vol. 6, no. 12, pp. 1621–1633, December 1997.
- [4] A. Patti, M. Sezan, and A. Tekalp, “High-resolution image reconstruction from a low-resolution image sequence in the presence of time-varying motion blur,” in *Proceedings of the IEEE International Conference on Image Processing*, Austin, TX, vol. 1, 1994, pp. 343–347.
- [5] R. C. Hardie, K. J. Barnard, J. G. Bogner, E. E. Armstrong, and E. A. Watson, “High resolution image reconstruction from a sequence of rotated and translated frames and its application to an infrared imaging system,” *Optical Engineering*, vol. 37, no. 1, pp. 247–260, January 1998.
- [6] M. S. Alam, J. G. Bogner, R. C. Hardie, and B. J. Yasuda, “Infrared image registration and high-resolution reconstruction using multiple translationally shifted aliased video frames,” *IEEE Transactions on Instrumentation and Measurement*, vol. 49, no. 5, pp. 923–915, October 2000.
- [7] R. Y. Tsai and T. S. Huang, “Multiframe image restoration and registration,” in *Advances in Computer Vision and Image Processing*:

- Image Reconstruction from Incomplete Observations, T. S. Huang, Ed., vol. 1. London: JAI Press, 1984, pp. 317–339.
- [8] N. K. Bose, H. C. Kim, and H. M. Valenzuela, “Recursive total least squares algorithm for image reconstruction from noisy undersampled frames,” *Multidimensional Systems and Signal Processing*, vol. 4, no. 3, pp. 253–268, July 1993.
 - [9] J. Yang, J. Wright, T. Huang, and Yi Ma. Image super-resolution via sparse representation. *IEEE Transactions on Image Processing (TIP)*, vol. 19, issue 11, 2010.
 - [10] Martin A. Fischler and Robert C. Bolles (June 1981). “Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography”. *Comm. of the ACM* 24 (6): 381–395
 - [11] Jordan Ellenberg, ”Fill in the Blanks: Using Math to Turn Lo-Res Datasets Into Hi-Res Samples”, 22 February 2010, *Wired Magazine* http://www.wired.com/2010/02/ff_algorithm/
 - [12] D.L. Donoho and X. Huo, “Uncertainty principles and ideal atomic decomposition,” *IEEE Trans. Inform. Theory*, vol. 47, no. 7, pp. 2845–2862, Nov. 2001.
 - [13] E. Candes, “Compressive sensing,” in *Proceedings of the International Congress of Mathematicians*, vol. 3, pp. 1433–1452, 2006.
 - [14] D. L. Donoho, “Compressed sensing,” *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.
 - [15] John G. Proakis, Dimitris G. Manolakis , “Digital Signal Processing - Principles, Algorithms, Implementations” 4th edition (Pearson International Edition), Pearson Education, Chapter 1.4.2: The Sampling Theorem
 - [16] Rafael C. Gonzalez, Richard E. Woods , “Digital Image Processing” 3rd edition (Pearson International Edition), Pearson Education, Chapter 1.4.2: The 2D Discrete Fourier Transform and its Inverse
 - [17] S. Farsiu, M. Elad, P. Milanfar, “Fast and Robust Multiframe Super Resolution” *IEEE Transactions on Image Processing*, vol. 13, no. 10, October 2004
 - [18] ”Image Registration”, in *Wikipedia: The Free Encyclopedia*; (Wikimedia Foundation Inc., updated 2 December 2014, 15:50 UTC) http://en.wikipedia.org/wiki/Image_registration
 - [19] M. Davenport, M. Duarte, Y. Eldar, G. Kutyniok, ”Introduction to Compressed Sensing”, Stanford University - Department of Statistics
 - [20] F. Bach, R. Jenatton, J. Mairal, G. Obozinski, ”Convex Optimization with Sparsity-Inducing Norms”, Institut National de Recherche en Informatique et en Automatique (INRIA)
 - [21] H. Rauhut, K. Schnass, and P. Vandergheynst, “Compressed sensing

- and redundant dictionaries,” IEEE Transactions on Information Theory, vol. 54, no. 5, May 2008.
- [22] M. Elad and M. Aharon, “Image denoising via sparse and redundant representations over learned dictionaries,” IEEE Transactions on Image Processing, vol. 15, pp. 3736–3745, 2006.
- [23] J. Mairal, G. Sapiro, and M. Elad, “Learning multiscale sparse representations for image and video restoration,” Multiscale Modeling and Simulation, vol. 7, pp. 214–241, 2008.
- [24] M. Aharon, M. Elad, and A. Bruckstein, “K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation,” IEEE Transactions on Signal Processing, vol. 54, no. 11, pp. 4311–4322, Nov. 2006.
- [25] H. Lee, A. Battle, R. Raina, and A. Y. Ng, “Efficient sparse coding algorithms,” in Advances in Neural Information Processing Systems (NIPS), pp. 801–808, 2007.
- [26] B. Olshausen and D. Field, “Sparse coding with an overcomplete basis set: A strategy employed by V1?,” Vision Research, vol. 37, no. 23, pp. 3311–3325, 1997.
- [27] “NP-hard”, in Wikipedia: The Free Encyclopedia; (Wikimedia Foundation Inc., updated 9 April 2015, 15:30 UTC) <http://en.wikipedia.org/wiki/NP-hard>
- [28] D. L. Donoho, “For most large underdetermined systems of linear equations, the minimal l_1 -norm solution is also the sparsest solution,” Communications on Pure and Applied Mathematics, vol. 59, no. 6, pp. 797–829, 2006.
- [29] “For most large underdetermined systems of linear equations, the minimal l_1 -norm near-solution approximates the sparsest near-solution,” Communications on Pure and Applied Mathematics, vol. 59, no. 7, pp. 907–934, 2006.
- [30] R. Tibshirani, “Regression shrinkage and selection via the lasso,” Journal of Royal Statistical Society, Series B, vol. 58, no. 1, 1996.
- [31] Jianchao Yang, Zhaowen Wang, Zhe Lin, and Thomas Huang. Coupled dictionary training for image super-resolution. IEEE Transactions on Image Processing (TIP), vol. 21, issue 8, pages 3467-3478, 2012.
- [32] W. T. Freeman, T. R. Jones, and E. C. Pasztor, “Example-based superresolution,” IEEE Computer Graphics and Applications, vol. 22, pp. 56–65, 2002.
- [33] W. T. Freeman, E. C. Pasztor, and O. T. Carmichael, “Learning low-level vision,” International Journal of Computer Vision, vol. 40, no. 1, pp. 25–47, 2000.
- [34] H. Chang, D.-Y. Yeung, and Y. Xiong, “Super-resolution through neighbor embedding,” in IEEE Conference on Computer Vision and

- Pattern Classification (CVPR), vol. 1, pp. 275–282, 2004.
- [35] J. Sun, N. N. Zheng, H. Tao, and H. Shum, “Image hallucination with primal sketch priors,” in IEEE Conference on Computer Vision and Pattern Recognition (CVPR), vol. 2, 2003, pp. 729–736.
- [36] imfill, Image Processing Toolbox Documentation, MathWorks
<http://www.mathworks.com/help/images/ref/imfill.html>
- [37] imresize, Image Processing Toolbox Documentation, MathWorks
<http://www.mathworks.com/help/images/ref/imresize.html>
- [38] A. Cohen, W. Dahmen, R. DeVore, “Compressed sensing and best k-term approximation”, in the Journal of the American Mathematical Society, vol. 22, pp. 211-231 , 2009.

