

Όραση Υπολογιστών

Εργασία 1η: Αναγνώριση περιοχών κειμένου του
εγγράφου & αποθορυβοποίηση

Γραμμένος-Γεώργιος Πολυμερίδης

A.M.: 58105

11 Νοεμβρίου 2025

Περιεχόμενα

1 Πρόβλημα	3
1.1 Περιγραφή Προβλήματος	3
2 Αποθορυβοποίηση	3
2.1 Είδη Θορύβου	3
2.1.1 Gaussian Noise	3
2.1.2 Salt & Pepper Noise	3
2.2 Τεχνικές	3
2.2.1 Gaussian Filtering	3
2.2.2 Median Filtering	5
3 Ανίχνευση Υποπεριοχών	9
3.1 Περίληψη Μεθόδου	9
3.2 Μέθοδος Integral	9
3.3 Επιφάνεια της υποπεριοχής που καταλαμβάνεται από κείμενο	9
4 Καταμέτρηση Λέξεων	10
5 Projection	14
5.1 Περίληψη Μεθόδου	14
5.2 Υλοποίηση	14
5.2.1 Εντοπισμός Υποπεριοχών	14
5.2.2 Εντοπισμός και Καταμέτρηση λέξεων	14
5.3 Συμπεράσματα	14
5.4 Γραφήματα – Projection	15
5.5 Αποτελέσματα	15
6 Κώδικας	16
7 Δομή Project	19
7.1 Πως να τρέξετε τα αποτελέσματα	19
8 Αποτελέσματα	20

1 Πρόβλημα

1.1 Περιγραφή Προβλήματος

Το πρόβλημα που έχουμε να αντιμετωπίσουμε είναι η αναγνώριση υποπεριοχών σε ένα κείμενο από σελίδα βιβλίου. Προκειμένου να γίνεται αξιόπιστα θα κάνουμε κάποιες παραδοχές.

- Δεν γνωρίζουμε την φωτεινότητα ή το πόσο θολή ή καθαρή είναι η εικόνα (Gaussian blur θόρυβος).
- Υπάρχει θόρυβος (Salt & Paper).
- Τα βιβλία είναι τυποποιημένα, ára υπάρχουν συγκεκριμένα μεγέθη γραμμών και γραμματοσειρών που μένουν σταθερά σε μέγεθος και διαστάσεις.

2 Αποθορυβοποίηση

2.1 Είδη Θορύβου

2.1.1 Gaussian Noise

Dummy

2.1.2 Salt & Pepper Noise

Dummy

2.2 Τεχνικές

2.2.1 Gaussian Filtering

Gaussian filtering

$$G_{ij} = \frac{1}{2\pi\sigma^2} e^{-\frac{(i-k)^2 + (j-k)^2}{2\sigma^2}}$$

Αυτού του είδους το φίλτρο θεωρείται μη-γραμμικό και θεωρείται μη αντιστρέψιμο. Είναι εφικτό ωστόσο με μερικές προσεγγίσεις να πάρουμε ένα καλύτερο αποτέλεσμα αλλά μπορούμε να θεωρήσουμε ότι δεν αντιστρέφεται για χάρην απλότητας στην συγκεκριμένη εφαρμογή και μονάχα να παρατηρήσουμε πως επηρεάζει στο τελικό αποτέλεσμα.

Π.χ.

$$\begin{pmatrix} 0.075 & 0.124 & 0.075 \\ 0.124 & 0.204 & 0.124 \\ 0.075 & 0.124 & 0.075 \end{pmatrix}$$

```
23 def gaussian_filter(image, kernel_size=5, sigma=0):
24     # Δημιουργία 2D Gaussian kernel
25     center = kernel_size // 2
26     kernel = np.zeros((kernel_size, kernel_size))
27     for i in range(kernel_size):
28         for j in range(kernel_size):
```

```

29     x = i - center
30     y = j - center
31     kernel[i, j] = np.exp(-(x ** 2 + y ** 2) / (2 * sigma ** 2))
32 kernel = kernel / kernel.sum()
33 # Συνέλιξη
34 pad = kernel_size // 2
35 padded = np.pad(image, pad, mode='edge')
36 output = np.zeros_like(image, dtype=np.float64)
37 for i in range(image.shape[0]):
38     for j in range(image.shape[1]):
39         region = padded[i:i + kernel_size, j:j + kernel_size]
40         output[i, j] = np.sum(region * kernel)
41 return output.astype(np.uint8)

```

Region 1
Implemented to improve directivity and gain, which are critical for L&S. The development of phased-array antennas for radar in the 1950s was a significant milestone [2]. In this era, a new initiative was undertaken to develop antenna arrays without moving the physical antennas, thereby greatly enhancing the speed and flexibility of L&S. With the advancement of antenna array technologies, array signal processing gradually became an active research area since the 1980s [3]. This period also introduced the windowed and sliding L&S methods based on array signal processing, including the famous MUSIC (Multiple Signal Classification), parametric maximum likelihood estimator, and ESPRIT (Estimation of Signal Parameters via Rotational Invariance Technique) methods. These methods laid an important foundation for array signal processing techniques for L&S.

The trend of the century brought about a digital revolution, further transforming the capabilities of antenna arrays. The integration of digital signal processing with antenna array technology led to the development of adaptive antenna systems that are capable of dynamically adjusting their patterns to optimize L&S accuracy. This era also witnessed the advent of multiple-input, multiple-output (MIMO) systems [4], which made it possible to increase the number of antennas to further improve communication performance through beamforming, spatial multiplexing, and diversity. Building on these advancements, the concept of MIMO radar emerged in the early 2000s [5], allowing each antenna to transmit different waveforms and thus providing more adaptive array configuration and higher L&S resolution.

The electromagnetic (EM) theory of near-field L&S states that antennas or antenna arrays can be divided into two regions: the far-field region and the near-field region (Fresnel-diffraction region [6], [7]). The demarcation between these regions is traditionally defined by the phase error-based Fraunhofer distance (also known as the Rayleigh distance), which is given by

$$d_F = \frac{\lambda D}{4}$$

Region 2
Region 2 denotes the aperture of the antenna array, and λ is the signal wavelength. The exact boundary between these regions can be defined differently depending on the application. Regardless, when the emitted signals propagate into the far-field region, its intensity is approximately a point source, which implies that the receiving antenna can only resolve the direction of the signal. However, the accurate spherical-wavefront model should be used to characterize the signal propagation within the near-field region, where the curvature of distance is considered. This property allows the antenna array to detect both the amplitude and the propagation direction of the signal. Research on the L&S of sources or targets in the near-field region dates back to as early as the 1980s [8], [9]. However, this research topic did not attract widespread interest due to the fact that the receiving antennas and signal frequencies were generally limited, rendering the sources or targets normally located in the far-field region.

Region 3
The landscape shifted with the introduction of the massive MIMO concept by Marzetta in 2010 [10]. The core idea of massive MIMO is to equip base stations with numerous antennas that are connected to fully digital transceiver blocks, allowing them to generate more simultaneous beams with high controllability. Massive MIMO was developed for communications but builds on enhancing the spatial resolution compared to smaller arrays, which also significantly improves L&S accuracy. These developments have made the massive MIMO system a fundamental building block of 5G networks and typically operate in the 3.5-GHz band. Since the emergence of 5G massive MIMO, researchers have continued to emphasize using even larger arrays with various form factors and characteristics, thereby increasing the physical size of the antenna arrays. Coupled with the trend toward exploiting high-frequency bands (centimeter and millimeter waves and beyond), which have very short wavelengths, the near-field region of antenna arrays may be significantly extended. For instance, an antenna array with a 1-m aperture operating at 30 GHz has a near-field region extending up to 200 m, making it highly likely for targets to be located within this near-field region. Against the alternative, the trend toward low-frequency bands has been devoted to the field of near-field L&S, focusing on various aspects such as signal modeling [6], [7], [11], [12], signal processing [13], [14], [15], [16], [17], [18], [19], and performance-based channel estimation [20]. This article aims to provide a tutorial review on near-field L&S from signal modeling to signal processing. Our goal is to highlight the most significant changes from traditional far-field L&S and elucidate the key intuitive ideas for leveraging the advantages of near-field L&S and addressing the associated challenges.

Region 4
Near-field properties
In this section, we will outline near-field propagation from the perspective of EM theory and discuss its relevance to the near-field L&S considered in this article. Let us consider a monochromatic wave with frequency f_c transmitted from a point source. The Fourier representation of the electric field observed at a distance r from the source, in any direction perpendicular to the propagation direction, is proportional to (12)

$$E(r) = \frac{jk}{2\pi r} \left(1 + \frac{j\lambda}{2\pi r} \frac{r^2}{(2\pi r)^2} \right) \quad (12)$$

where k denotes the intrinsic impedance, $\lambda = c/f_c$ denotes the wavelength with c being the speed of light, and $j = \sqrt{-1}$. The last two terms, which decay faster than $1/r$, diminish rapidly as r increases and are related to the radiation loss. These terms are associated with "near-field" propagation in EM theory, but they are more commonly referred to as *reactive near-field* in recent literature on near-field communications and L&S. It has been demonstrated that the reactive near-field region is much smaller than the wavelength of the source, even for an infinitely large antenna (or array) [22]. Therefore, in practical system design, only the radiating field

Region 5
The landscape shifted with the introduction of the massive MIMO concept by Marzetta in 2010 [10]. The core idea of massive MIMO is to equip base stations with numerous antennas that are connected to fully digital transceiver blocks, allowing them to generate more simultaneous beams with high controllability. Massive MIMO was developed for communications but builds on enhancing the spatial resolution compared to smaller arrays, which also significantly improves L&S accuracy. These developments have made the massive MIMO system a fundamental building block of 5G networks and typically operate in the 3.5-GHz band. Since the emergence of 5G massive MIMO, researchers have continued to emphasize using even larger arrays with various form factors and characteristics, thereby increasing the physical size of the antenna arrays. Coupled with the trend toward exploiting high-frequency bands (centimeter and millimeter waves and beyond), which have very short wavelengths, the near-field region of antenna arrays may be significantly extended. For instance, an antenna array with a 1-m aperture operating at 30 GHz has a near-field region extending up to 200 m, making it highly likely for targets to be located within this near-field region. Against the alternative, the trend toward low-frequency bands has been devoted to the field of near-field L&S, focusing on various aspects such as signal modeling [6], [7], [11], [12], signal processing [13], [14], [15], [16], [17], [18], [19], and performance-based channel estimation [20]. This article aims to provide a tutorial review on near-field L&S from signal modeling to signal processing. Our goal is to highlight the most significant changes from traditional far-field L&S and elucidate the key intuitive ideas for leveraging the advantages of near-field L&S and addressing the associated challenges.

Region 6
The landscape shifted with the introduction of the massive MIMO concept by Marzetta in 2010 [10]. The core idea of massive MIMO is to equip base stations with numerous antennas that are connected to fully digital transceiver blocks, allowing them to generate more simultaneous beams with high controllability. Massive MIMO was developed for communications but builds on enhancing the spatial resolution compared to smaller arrays, which also significantly improves L&S accuracy. These developments have made the massive MIMO system a fundamental building block of 5G networks and typically operate in the 3.5-GHz band. Since the emergence of 5G massive MIMO, researchers have continued to emphasize using even larger arrays with various form factors and characteristics, thereby increasing the physical size of the antenna arrays. Coupled with the trend toward exploiting high-frequency bands (centimeter and millimeter waves and beyond), which have very short wavelengths, the near-field region of antenna arrays may be significantly extended. For instance, an antenna array with a 1-m aperture operating at 30 GHz has a near-field region extending up to 200 m, making it highly likely for targets to be located within this near-field region. Against the alternative, the trend toward low-frequency bands has been devoted to the field of near-field L&S, focusing on various aspects such as signal modeling [6], [7], [11], [12], signal processing [13], [14], [15], [16], [17], [18], [19], and performance-based channel estimation [20]. This article aims to provide a tutorial review on near-field L&S from signal modeling to signal processing. Our goal is to highlight the most significant changes from traditional far-field L&S and elucidate the key intuitive ideas for leveraging the advantages of near-field L&S and addressing the associated challenges.

Region 7
Near-field properties
In this section, we will outline near-field propagation from the perspective of EM theory and discuss its relevance to the near-field L&S considered in this article. Let us consider a monochromatic wave with frequency f_c transmitted from a point source. The Fourier representation of the electric field observed at a distance r from the source, in any direction perpendicular to the propagation direction, is given by

$$E(r) = \frac{jk}{2\pi r} \left(1 + \frac{j\lambda}{2\pi r} \frac{r^2}{(2\pi r)^2} \right) \quad (1)$$

where D denotes the aperture of the antenna (array), and λ is the signal wavelength. The exact boundary between these regions can be defined differently depending on the application. Regardless, when the emitted signal propagates into the far-field region, its waveform is approximated as plane, which implies that it receives no information about the radiation pattern of the source. However, the accurate spherical-wavefront model should be used to characterize the signal propagation within the near-field region, where the curvature of distance r is considered. This property allows the antenna array to detect both the amplitude and the propagation direction of the signal. Research on the L&S of sources or targets in the near-field region dates back to as early as the 1980s [8], [9].

However, this research topic did not attract widespread interest due to the fact that the receiving antennas and signal frequencies were generally limited, rendering the sources or targets normally located in the far-field region. Therefore, in practical system design, only the radiating field

Region 8
IEEE SIGNAL PROCESSING MAGAZINE | January 2025

Region 9
IEEE SIGNAL PROCESSING MAGAZINE | January 2025

Region 5
IEEE SIGNAL PROCESSING MAGAZINE | January 2025

Region 6
IEEE SIGNAL PROCESSING MAGAZINE | January 2025

(α') Median kernel 5×5 και $\sigma = 0$

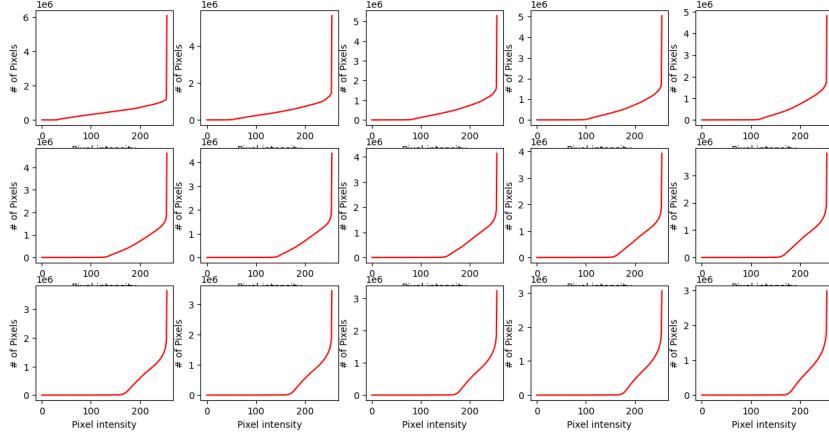
(β') Median kernel 31×31 και $\sigma = 0$

Σχήμα 1: Σύγκριση δύο διαφορετικών τιμών για φίλτρο Gaussian

Παρατηρούμε ότι όσο αυξάνεται το kernel size τόσο πιο εύκολα συγχωνεύονται μεγάλα κενά διαστήματα, με αποτέλεσμα να χάνουμε την ακρίβεια στην αναγνώριση και αριθμηση των υποτεριοχών.

for L&S. The develop
r in the 1950s was a sig
initiative steered the d
out moving the physica
he speed and flexibility
ma.array technologies.
for L&S. The develop
r in the 1950s was a sig
initiative steered the d
out moving the physica
he speed and flexibility
ma.array technologies.
Kernel 13
for L&S. The develop
r in the 1950s was a sig
initiative steered the d
out moving the physica
he speed and flexibility
ma.array technologies.
for L&S. The develop
r in the 1950s was a sig
initiative steered the d
out moving the physica
he speed and flexibility
ma.array technologies.
Kernel 15
for L&S. The develop
r in the 1950s was a sig
initiative steered the d
out moving the physica
he speed and flexibility
ma.array technologies.
for L&S. The develop
r in the 1950s was a sig
initiative steered the d
out moving the physica
he speed and flexibility
ma.array technologies.
Kernel 17
for L&S. The develop
r in the 1950s was a sig
initiative steered the d
out moving the physica
he speed and flexibility
ma.array technologies.
for L&S. The develop
r in the 1950s was a sig
initiative steered the d
out moving the physica
he speed and flexibility
ma.array technologies.
Kernel 19
for L&S. The develop
r in the 1950s was a sig
initiative steered the d
out moving the physica
he speed and flexibility
ma.array technologies.
for L&S. The develop
r in the 1950s was a sig
initiative steered the d
out moving the physica
he speed and flexibility
ma.array technologies.
Kernel 21
for L&S. The develop
r in the 1950s was a sig
initiative steered the d
out moving the physica
he speed and flexibility
ma.array technologies.
for L&S. The develop
r in the 1950s was a sig
initiative steered the d
out moving the physica
he speed and flexibility
ma.array technologies.
Kernel 23
for L&S. The develop
r in the 1950s was a sig
initiative steered the d
out moving the physica
he speed and flexibility
ma.array technologies.
for L&S. The develop
r in the 1950s was a sig
initiative steered the d
out moving the physica
he speed and flexibility
ma.array technologies.
Kernel 25
for L&S. The develop
r in the 1950s was a sig
initiative steered the d
out moving the physica
he speed and flexibility
ma.array technologies.
for L&S. The develop
r in the 1950s was a sig
initiative steered the d
out moving the physica
he speed and flexibility
ma.array technologies.
Kernel 27
for L&S. The develop
r in the 1950s was a sig
initiative steered the d
out moving the physica
he speed and flexibility
ma.array technologies.
for L&S. The develop
r in the 1950s was a sig
initiative steered the d
out moving the physica
he speed and flexibility
ma.array technologies.
Kernel 29
for L&S. The develop
r in the 1950s was a sig
initiative steered the d
out moving the physica
he speed and flexibility
ma.array technologies.
for L&S. The develop
r in the 1950s was a sig
initiative steered the d
out moving the physica
he speed and flexibility
ma.array technologies.
Kernel 31

(a') Δοκιμή διάφορων Gaussian kernels

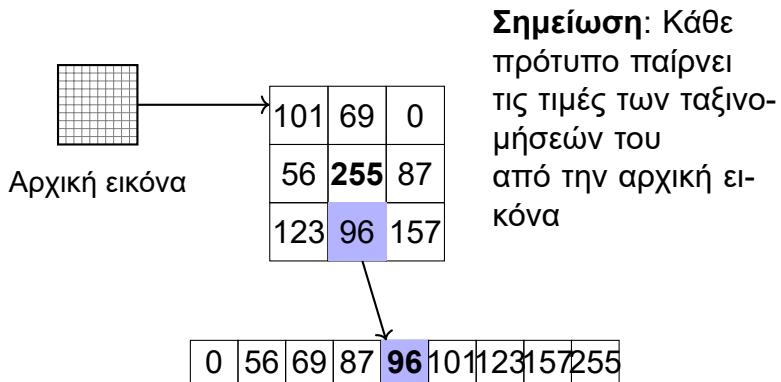


Σχήμα 2: Δοκιμή Guassian φίλτρου για διάφορες τιμές kernel size με $\sigma = 0$

2.2.2 Median Filtering

Ο τρόπος που λειτουργεί το median filtering είναι μη γραμμικός και αυτό του δίνει κιόλας το πλεονέκτημα έναντι γραμμικών φίλτρων στη μείωση του θορύβου, επίσης το median filter αναφέρεται και ως salt & pepper filter γιατί έχει την μικρότερη παραμόρφωση για αυτή την εφαρμογή

$$\hat{f}(x, y) = \underset{(s,t) \in S_{xy}}{\text{median}}\{g(s, t)\}$$



cial for L&S. The development of phased-array antennas for radar in the 1950s was a significant milestone [2]. In this era, a

sive MIMO concept by Marzetta in 2010 [10]. The core idea of massive MIMO is to equip base stations with numerous anten-

$$(\alpha') 1 \times 1$$

cial for L&S. The development of phased-array antennas for radar in the 1950s was a significant milestone [2]. In this era, a

sive MIMO concept by Marzetta in 2010 [10]. The core idea of massive MIMO is to equip base stations with numerous anten-

$$(\beta') 3 \times 3$$

cial for L&S. The development of phased-array antennas for radar in the 1950s was a significant milestone [2]. In this era, a

sive MIMO concept by Marzetta in 2010 [10]. The core idea of massive MIMO is to equip base stations with numerous anten-

$$(\gamma') 5 \times 5$$

cial for L&S. The development of phased-array antennas for radar in the 1950s was a significant milestone [2]. In this era, a

sive MIMO concept by Marzetta in 2010 [10]. The core idea of massive MIMO is to equip base stations with numerous anten-

$$(\delta') 7 \times 7$$

cial for L&S. The development of phased-array antennas for radar in the 1950s was a significant milestone [2]. In this era, a

sive MIMO concept by Marzetta in 2010 [10]. The core idea of massive MIMO is to equip base stations with numerous anten-

$$(\epsilon') 9 \times 9$$

cial for L&S. The development of phased-array antennas for radar in the 1950s was a significant milestone [2]. In this era, a

sive MIMO concept by Marzetta in 2010 [10]. The core idea of massive MIMO is to equip base stations with numerous anten-

$$(\zeta') 11 \times 11$$

cial for L&S. The development of phased-array antennas for radar in the 1950s was a significant milestone [2]. In this era, a

sive MIMO concept by Marzetta in 2010 [10]. The core idea of massive MIMO is to equip base stations with numerous anten-

$$(\eta') 15 \times 15$$

cial for L&S. The development of phased-array antennas for radar in the 1950s was a significant milestone [2]. In this era, a

sive MIMO concept by Marzetta in 2010 [10]. The core idea of massive MIMO is to equip base stations with numerous anten-

$$(\theta') 17 \times 17$$

cial for L&S. The development of phased-array antennas for radar in the 1950s was a significant milestone [2]. In this era, a

sive MIMO concept by Marzetta in 2010 [10]. The core idea of massive MIMO is to equip base stations with numerous anten-

$$(\iota') 19 \times 19$$

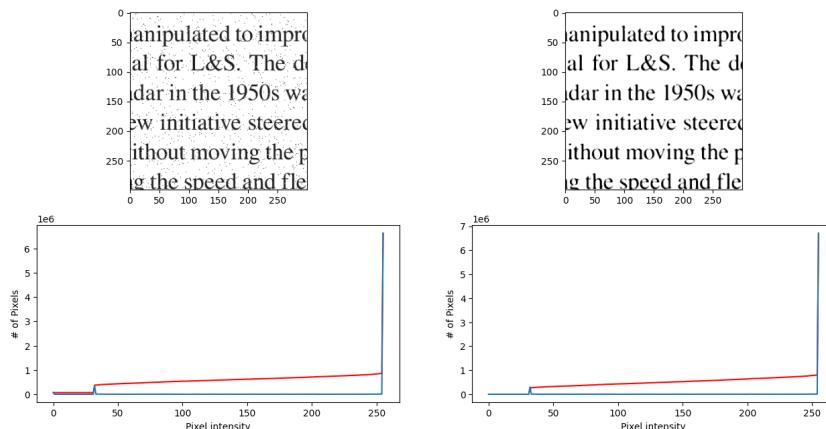
Σχήμα 3: Αντίκτυπο στις λεπτομέριες του εγγράφου Median φίλτρου για διάφορες τιμές kernel size

```

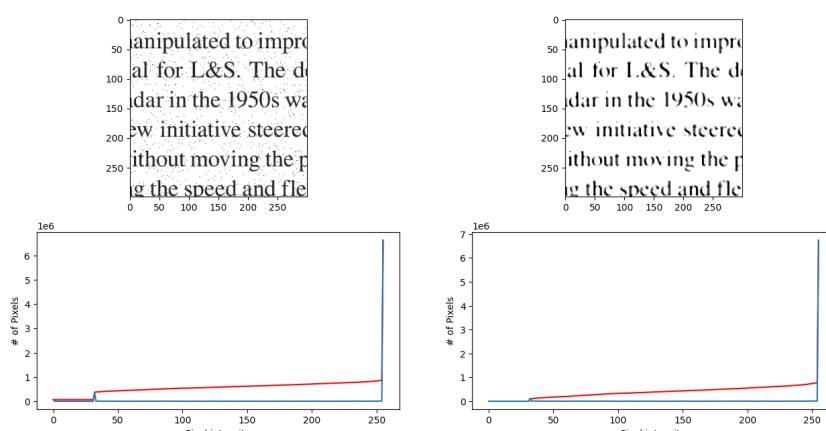
7 def median_filter(image, kernel_size=5):
8     h, w = image.shape
9     pad = kernel_size // 2
10    filtered = np.zeros_like(image)
11    padded = np.pad(image, pad, mode='edge')
12    for i in range(h):
13        for j in range(w):
14            window = padded[i:i + kernel_size, j:j + kernel_size]
15            filtered[i, j] = np.median(window)
16    return filtered

```

Το padding στις περισσότερες εφαρμογές με εικόνες το προσθέτω προκειμένου να είμαι σίγουρος ότι στις ακμές της εικόνας θα υπάρχει κάποια νοητή τιμή προκειμένου να είναι εφικτές οι πράξεις, διαφορετικά θα πρέπει να περικόψω την εικόνα στο μισό της διάστασης του παραθύρου μου, κάτι που προκαλεί απώλεια πληροφορίας και δεν το επιθυμούμε.



(α') Median kernel 3×3



(β') Median kernel 5×5

Σχήμα 4: Δοκιμή Median φίλτρου για διάφορες τιμές kernel size

for L&S. The development in the 1950s was a significant initiative steered the physical speed and flexibility in array technologies.

The A/D conversion in the 1970s was a major breakthrough in the 1970s was a significant initiative steering the physical speed and flexibility in array technologies.

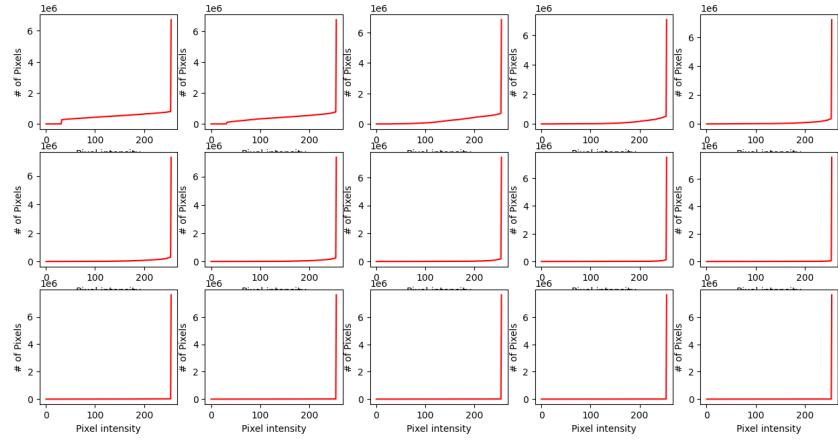
Kernel: 13 Kernel: 15 Kernel: 17 Kernel: 19 Kernel: 21 Kernel: 23 Kernel: 25 Kernel: 27 Kernel: 29 Kernel: 31

for L&S. The development in the 1950s was a significant initiative steering the physical speed and flexibility in array technologies.

The A/D conversion in the 1970s was a major breakthrough in the 1970s was a significant initiative steering the physical speed and flexibility in array technologies.

Kernel: 13 Kernel: 15 Kernel: 17 Kernel: 19 Kernel: 21 Kernel: 23 Kernel: 25 Kernel: 27 Kernel: 29 Kernel: 31

(α') Μεσημέτριος φίλτρος 3×3



(β') Ιστογράμματα

Σχήμα 5: Δοκιμή Median φίλτρου για διάφορες τιμές kernel size

3 Ανίχνευση Υποπεριοχών

3.1 Περίληψη Μεθόδου

Η πρόκληση είναι ότι δεν επιτρέπεται να χρησιμοποιήσω την μέθοδο για την εύρεση των ισουψών καμπύλων, βέβαια αυτός ο περιορισμός δεν έχει τόση σημασία διότι είναι εφικτό. Οι μορφωλογικές πράξεις με έναν αρκετά μεγάλο kernel αλλά όχι τόσο μεγάλο ώστε να ενώνει μεγάλα διαστήματα που αποτελείται από σειρές και παραγράφους μπορούν να δώσουν την συνολική επιφάνεια που καταλαμβάνεται από κείμενο.

```
23   text_integral = integral(text_mask)
24   (totalLabels, label_ids, values, centroid) =
25     → cv2.connectedComponentsWithStats(morph_close, 4)
26
27   detected = cv2.cvtColor(dut_array_image.copy(),cv2.COLOR_GRAY2BGR)
```

Η συνάρτηση που χρησιμοποιώ αντί για την cv2.findContours, είναι η cv2.connectedComponentsWithStats η οποία τρέχει τον αλγόριθμο Spaghetti4C [2] για connectivity. (Υπάρχει η επιλογή και με 8 γειτονικά σημεία ωστόσο εν προκειμένω χρησιμοποιώ τέσσερα)

3.2 Μέθοδος Integral

Εφόσον πρέπει οι υπολογισμοί να είναι ανεξάρτητοι του μεγέθους του παραθύρου προκειμένου να βρούμε την μέση διαβάθμιση του γκρί, αυτό μπορούμε να το πετύχουμε με το να υπολογίσουμε μια φορά το άθροισμα των προηγούμενων pixel και έπειτα επαναχρησιμοποιώντας τις τιμές αυτές με τρείς προσθέτες εισιτηρίους και μια αφαίρεση

```
4 def integral(image):
5   int = np.cumsum(image, axis=0, dtype=np.float32)
6   int = np.cumsum(int, axis=1, dtype=np.float32)
7   return np.pad(int, ((1,0),(1,0)),mode='constant')
```

Και εδώ προσθέτω padding, για να μην φύγω των ορίων των πινάκων. Η μέθοδος cumsum ουσιαστικά είναι cumulative sum [1], και αθροίζει και τα προηγούμενα, δηλαδή αυτό που χρειαζόμαστε στην περίπτωση μας.

3.3 Επιφάνεια της υποπεριοχής που καταλαμβάνεται από κείμενο

Εφόσον η σελίδα αποτελείται από κείμενο και κενά, θέλουμε να χρησιμοποιήσουμε ένα κατώφλι για να πάρουμε μόνο τα εικονοστοιχεία που είναι πράγματι κείμενο, χρησιμοποιούμε μάσκα για να απομονώσουμε το κείμενο.

```
20   text_mask = np.zeros_like(thresh)
21   text_mask[thresh == 0] = 255
22
23   text_integral = integral(text_mask)
24
25   sub_thresh = thresh[y1:(y1 + h1), x1:(x1 + w1)]
26   # Κάνω morphological close για να ενώσω τα γράμματα σε λέξεις
27   word_kernel = cv2.getStructuringElement(cv2.MORPH_RECT, (5, 5))
28   word_close = cv2.morphologyEx(sub_thresh, cv2.MORPH_CLOSE, word_kernel)
```

$p_{1,1}$	$p_{1,2}$	$p_{1,3}$
$p_{2,1}$	$p_{2,2}$	$p_{2,3}$
$p_{3,1}$	$p_{3,2}$	$p_{3,3}$

(α') Εικόνα Εισόδου I (με περιοχή Σ)

0	0	0	0
0	(D)	$S_{1,2}$	(C)
0	$S_{2,1}$	$S_{2,2}$	$S_{2,3}$
0	(B)	$S_{3,2}$	(A)

(β') Integral Image S (Υπολογισμός Σ)

Σχήμα 6: Σχηματική απεικόνιση υπολογισμού αθροίσματος περιοχής με τη μέθοδο Integral Image. Αριστερά (6α'): Η αρχική εικόνα I με την επιθυμητή περιοχή Σ (κόκκινο πλαίσιο). Δεξιά (6β'): Η αντίστοιχη Integral Image S . Το αθροίσμα της Σ υπολογίζεται άμεσα από τις 4 γωνιακές τιμές: $II_{22,33} = \mathbf{A} - \mathbf{B} - \mathbf{C} + \mathbf{D}$. Τα αθροίσματα integral image στο συγκεκριμένο παράδειγμα ισούνται με $A = II_{2,2} = p_{1,1} + p_{1,2} + p_{1,3}$. $B = II_{2,2} = p_{1,1} + p_{1,2} + p_{1,3} + p_{2,1} + p_{2,2} + p_{2,3} + p_{3,1}$. Γενικότερα $II_{A_x, A_y} = \sum_{i <= A_x, j <= A_y} p_{i,j}$.

4 Καταμέτρηση Λέξεων

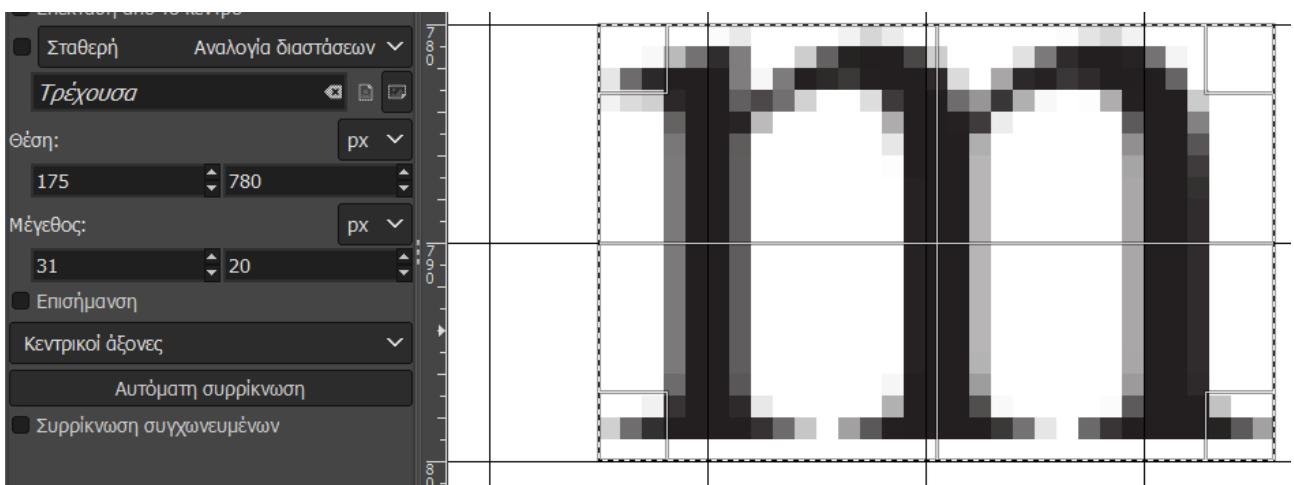
Η απλή προσέγγιση είναι εφόσον τα γράμματα είναι πολύ κοντά συνδεδεμένα αλλά τα διαστήματα μεταξύ των λέξεων μεγαλύτερα, τότε μέσω ενός μικρού πυρήνα (kernel) χρησιμοποιώντας $A \bullet B$ (Morph Close), τότε είναι εφικτό να διαχωρίσω τις λέξεις μεταξύ τους.

```

49 assert sub_area <= bounding_box_area
50
51 sub_thresh = thresh[y1:(y1 + h1), x1:(x1 + w1)]
52 # Κάνω morphological close για να ενώσω τα γράμματα σε λέξεις
53 word_kernel = cv2.getStructuringElement(cv2.MORPH_RECT, (5, 5))
54 word_close = cv2.morphologyEx(sub_thresh, cv2.MORPH_CLOSE, word_kernel)
55
56 # Μετράω τα connected components = λέξεις
57 word_count, _, sub_values, _ = cv2.connectedComponentsWithStats(word_close, 4)
58 # cv2.imwrite(f'./Report/text_segmentation/parts/{i}.png', word_close)
59
60 if DEBUG:
61     print(f'-----Region {i} -----')
62     print(f'Bounding box area(px): {bounding_box_area}')

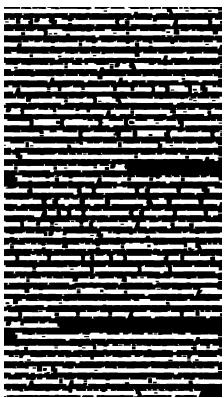
```

Στο συγκεκριμένο snippet χρησιμοποιώ ένα kernel 15×5 , αν υποθέσουμε για τυποποιημένους χαρακτήρες ότι είναι περίπου 20px επί 20px και για ένα διάστημα μεταξύ δύο λέξεων. Ένα αρκετά μικρό kernel ώστε να ενώνει αυτό το διάστημα είναι αρκτό, ασφαλώς το διάστημα μεταξύ γραμμάτων είναι πολύ μικρότερο από το πλήρε πλάτος των 20px, για την ακρίβεια είναι 1px από μέτρηση σε κατάλληλο πρόγραμμα επεξεργασίας εικόνας στο συγκεκριμένο παράδειγμα. Ενώ το πραγματικό μέγεθος δίνεται στο (7)

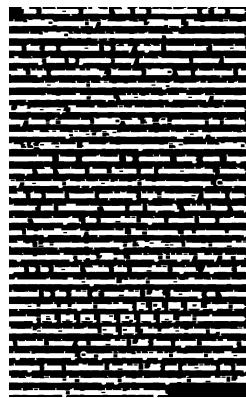


Σχήμα 7: Μέγεθος χαρακτήρα

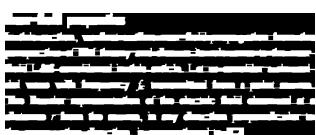
Παρατηρούμε ότι η (8β') ενώ αποτελείται από 6 λέξεις, ότι αυτό το kernel δεν είναι κατάλληλο γιατί ενώνει μορφολογικά λέξεις που είναι στην πραγματικότητα διαφορετικές οπότε παίρνουμε δύο λέξεις.



(α') Περιοχή 1



(γ') Περιοχή 3



(δ') Περιοχή 4



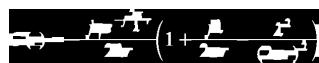
(ε') Περιοχή 5



(ζ') Περιοχή 6



(ζ') Περιοχή 7



(η') Περιοχή 8



(θ') Περιοχή 9



(ι') Περιοχή 10



(ια') Περιοχή 11



(ιβ') Περιοχή 12

Σχήμα 8: Διαχωρισμός Λέξεων: Προσπάθεια πρώτη 15×5

manipulated to improve directivity and gain, which are crucial for L&S. The development of phased-array antennas for radar in the 1950s was a significant milestone [2]. In this era, a new initiative altered the direction of the beam electronically without moving it physically, thereby greatly extending the potential flexibility of L&S. With the introduction of antenna array technologies, array signal processing has gradually become an active research area since the 1990s [3]. This period, extending into the 1990s, witnessed a surge of classical L&S methods for array signal processing, including the Minimum MSE (MMSE) (Minimum Mean Square Error), maximum-likelihood estimator, and ESPRIT (Estimation of Signal Parameters via Rotational Invariance Technique) methods. These methods laid an important foundation for array signal processing in L&S.

In the turn of the century, bright shows a digital revolution, further transforming the capabilities of antenna arrays. The integration of digital signal processing with antenna array techniques led to the development of adaptive antenna systems that are capable of dynamically adjusting the antenna pattern to the L&S scenario. This era also witnessed the birth of multiple-input, multiple-output (MIMO) systems [4], which use multiple antennas at both the transmitter and receiver to improve communication performance through beamforming, spatial multiplexing, and diversity. Building on these successes, the concept of MIMO radar was proposed in the early 2000s [5], allowing each antenna to transmit different waveforms and thus providing more adaptive array configuration and higher L&S resolution.

The electromagnetic (EM) radiation field emitted by antenna arrays can be divided into two regions: the near-field field (Prahladita region) and the radiative near-field (Prasad) region [6], [7]. The demarcation between these regions is traditionally defined by the phase error-based Prahlad distance, which is given by

(α') Περιοχή 1

(β') Περιοχή 2

(γ') Περιοχή 3

Near-field properties

In this section, we first outline near-field propagation from the perspective of EM theory and discuss its relevance to the near-field L&S considered in this article. Let us consider a sinusoidal wave with frequency f_s transmitted from a point source. The Fourier representation of the electric field observed at a distance r from the source, in any direction perpendicular to the propagation direction, is proportional to [12]

(δ') Περιοχή 4

$$d_{\text{FA}} = \frac{2D^2}{\lambda}$$

(ε') Περιοχή 5

where D denotes the aperture of the antenna (array), and λ is the signal wavelength. The exact boundary between these regions can be defined differently depending on the application. Regardless, when the emitted signal propagates into the far-field region, its wavefront is approximated as planar, which implies that the receiving antenna array can only resolve the direction of the signal. However, the accurate spherical-wavefront model should be used to characterize the signal propagation within the near-field region, where the curvature is distance dependent. This property allows the antenna array to determine both the direction and the propagation distance of the signal. Research on the L&S of sources or targets in the near-field region dates back to as early as the 1980s [8], [9]. However, this research topic did not attract widespread attention during that era. This is because array apertures and signal frequencies were generally limited, rendering the sources or targets normally located in the far-field region.

(ζ') Περιοχή 7

$$G(r) = -\frac{j\pi r^2 Y}{2\lambda r} \left(1 + \frac{j\lambda}{2\pi r} - \frac{\lambda^2}{(2\pi r)^2}\right)$$

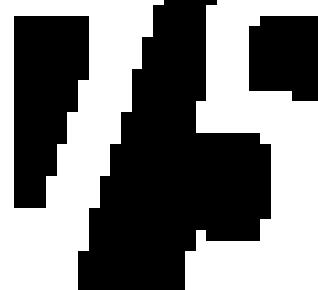
(η') Περιοχή 8

where η denotes the intrinsic impedance, $\lambda = c/f_s$ denotes the wavelength with c being the speed of light, and $j = \sqrt{-1}$. The last two terms, which decay faster than $1/r$, diminish rapidly as r increases and do not contribute to EM radiation. These terms are associated with "near-field" propagation in EM theory, but they are more commonly referred to as *reactive near-field* in recent literature on near-field communications and L&S. It has been demonstrated that the reactive near field is confined within a few wavelengths from the source, even for an infinitely large antenna (or array) [22]. Therefore, in practical system design, only the radiating field

(ι') Περιοχή 10

(ια') Περιοχή 11

(θ') Περιοχή 9



(ιβ') Περιοχή 12

Σχήμα 9: Διαχωρισμός Λέξεων: Προσπάθεια δεύτερη 5×5

—Region 12 —

Bounding box area(px): 870

Area (px): 810

Words : 3

Mean gray-level value in bounding box: 72722.96296296296

Πλέον έχουμε σίγουρα καλύτερο αποτέλεσμα παρόλα αυτά δεν πέρνουμε το βέλτιστο δυνατό αποτέλεσμα, όμως θα το αφήσουμε όπως έχει γιατί είναι πάρα πολύ κοντά προσεγγιστικά, επιπλέον στην μεγάλη την περιοχή (9α') λαμβάνουμε 313 αριθμός λέξεων, το οποίο είναι επίσης πάρα πολύ κοντά πτοιοτικά.

The landscape shifted with the introduction of the massive MIMO concept by Marzetta in 2010 [10]. The core idea of massive MIMO is to equip base stations with numerous antennas that are connected to fully digital transceiver branches, allowing them to generate many simultaneous narrow beams with high controllability. Massive MIMO was developed for communications but builds on enhancing the spatial resolution of the signals, which is the signature advantage over L&S systems. These forward-thinking researchers also defined a key form of massive MIMO with 32–64 antennas the fundamental building block of 5G networks and typically operate in the 3.5-GHz band. Since the emergence of 5G massive MIMO, researchers have continued to envision using even larger arrays with various form factors and characteristics, inevitably increasing the physical size of the antenna arrays. Coupled with the interest toward exploring the opportunities presented by the massive MIMO era, Sengupta, which have been for about wavelength, the near-field region around base stations can be significantly expanded. For instance, an antenna array with a 1-m aperture operating at 30 GHz can have a near-field region extending up to 200 m, making it highly likely for targets to be located within this near-field region. Against the aforementioned background, a lot of research efforts have been devoted to the field of near-field L&S, and various valuable contributions have been made [6], [7], [11], [12], [13], [14], [15], [16], [17], [18], [19], and performance-bound characterization [20], [21]. This article aims to provide a tutorial review on near-field L&S from signal modeling to signal processing. Our goal is to highlight the most significant changes from traditional far-field L&S and elucidate the key intuitive ideas for leveraging the advantages of near-field L&S and addressing the associated challenges.

5 Projection

5.1 Περίληψη Μεθόδου

Η μέθοδος αυτή στηρίζεται στα ιστογράμματα του οριζόντιου και κάθετου προφίλ ουσιαστικά έτσι μπορούμε να δούμε που υπάρχουν γραμμές και που στήλες, οπότε οι κοινές τομές θα μας δίνουν την πληροφορία που αρχίζει και που τελειώνει το κάθε segment. Η μέθοδος αυτή εξαρτάται από την πυκνότητα των σκουρόχρωμων γραμμάτων / περιοχών, η απλή μέθοδος αυτό το επιτυγχάνει ορίζοντας ένα κατώφλι σε σχέση με το απόλυτο μέγιστο όλων των τιμών.

Σαφώς αυτό είναι και το πρόβλημα, επειδή μπορεί να υπάρχει μεγάλη διακύμανση μεταξύ ολικού και τοπικών μέγιστων, σε αυτή την περίπτωση εξετάζουμε την απλή αλλά βελτιωμένη μέθοδο projection profiling.

5.2 Υλοποίηση

5.2.1 Εντοπισμός Υποπεριοχών

Η προσέγγιση που επιλέγω για να είναι αξιόπιστη είναι να διαλέξω τις περιοχές που έχουν τις μικρότερες τιμές, δηλαδή είναι κενές και να την αφαιρέσω από την αρχική δυαδική εικόνα.

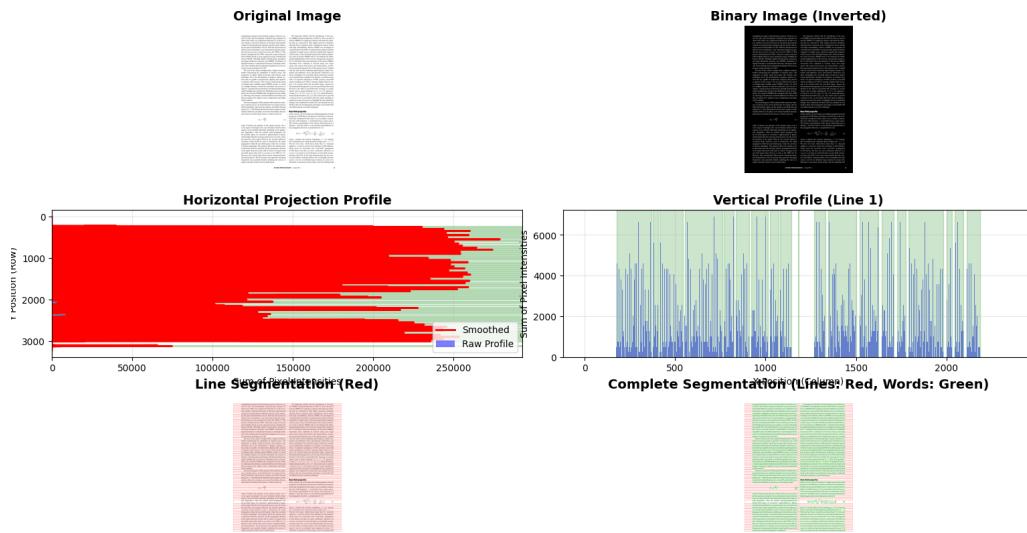
5.2.2 Εντοπισμός και Καταμέτρηση λέξεων

Αυτό μπορεί να γίνει και μόνο με horizontal profiling, έχοντας πρώτη γνώση τις διαστάσεις του κενού

5.3 Συμπεράσματα

Η μέθοδος αυτή είναι πολύ πιο αξιόπιστη από την προηγούμενη, που πρέπει να έχουμε πρότερη γνώση για το μέγεθος των γραμματοσειρών, το κενό των διαστημάτων.

5.4 Γραφήματα – Projection



Σχήμα 10: Τελικό αποτέλεσμα

5.5 Αποτελέσματα

manipulated to improve directivity and gain, which are crucial for L&S. The development of phased-array antennas for radar in the 1950s was a significant milestone [2]. In this era, a major breakthrough was the introduction of the parabolic dish, without moving the physical antenna, thereby greatly enhancing the speed and flexibility of L&S. With the advancement of technology, the development of MIMO has also made L&S truly become an active research area since the 1980s [3]. This paper aims to introduce the basic concepts of L&S and the L&S methods based on wave signal processing, including the famous MUSIC (Multiple Signal Classification), parametric MUSIC, MMSE (Minimum Mean Square Error), and Signal Parameters via Rotational Invariance Technique methods. These methods laid an important foundation for array signal processing.

The turn of the century brought about a digital revolution in L&S. The development of digital signal processing and the integration of digital signal processing with antenna array technology led to the development of adaptive antenna systems, which can automatically adjust the beamforming weights to optimize L&S accuracy. This era also witnessed the advent of multiple-input, multiple-output (MIMO) systems [4], which have greatly improved the performance of L&S. In order to improve communication performance through beamforming, the concept of MIMO radar emerged in the early 2000s [5], allowing each antenna to transmit different waveforms and receive different adaptive array configurations and higher L&S resolution.

In a document classification system, the collected omnidirectional or directional signals can be divided into two regions: the near-field (Fraunhofer region) and the radiative near-field (Fresnel region). The boundary between these two regions is usually defined by the plane-wave-based Fraunhofer distance (also known as the Rayleigh distance), which is given by

$$D = \frac{\lambda}{2\pi} \sqrt{r^2 + \left(\frac{\lambda}{2}\right)^2}$$

where D denotes the aperture of the antenna array, and λ is the signal wavelength. The main boundary between these regions can be defined differently depending on the application. For example, when the emitted signal propagates into the free space, the boundary is given by

$$d = \frac{\lambda}{2\pi} \sqrt{r^2 + \left(\frac{\lambda}{2}\right)^2}$$

which implies that the receiving antenna array can only resolve the source when it is located within the distance d .

A waveform model should be used to characterize the signal propagation within the near-field region, where the curvature of the wavefront is dominant. The wavefront curvature is used to determine both the direction and the propagation distance of the signal. Research on the L&S of sources or targets in the near-field region has been conducted for many years [6–10]. However, this research topic did not attract widespread attention until the late 1990s, when the interest in L&S increased. At that time, the available signal processing techniques were still limited, and the frequencies were generally limited, rendering the sources or targets normally located in the far-field region.

In addition to the limitations in signal processing, the near-field region is also characterized by the limitation of the near-field L&S. The development of phased-array antennas for radar in the 1950s was a significant milestone [11]. The introduction of the parabolic dish, without moving the physical antenna, thereby greatly enhancing the speed and flexibility of L&S. With the advancement of technology, the development of MIMO has also made L&S truly become an active research area since the 1980s [12]. This paper aims to introduce the basic concepts of L&S and the L&S methods based on wave signal processing, including the famous MUSIC (Multiple Signal Classification), parametric MUSIC, MMSE (Minimum Mean Square Error), and Signal Parameters via Rotational Invariance Technique methods. These methods laid an important foundation for array signal processing.

The turn of the century brought about a digital revolution in L&S. The development of digital signal processing and the integration of digital signal processing with antenna array technology led to the development of adaptive antenna systems, which can automatically adjust the beamforming weights to optimize L&S accuracy. This era also witnessed the advent of multiple-input, multiple-output (MIMO) systems [13], which have greatly improved the performance of L&S. In order to improve communication performance through beamforming, the concept of MIMO radar emerged in the early 2000s [14], allowing each antenna to transmit different waveforms and receive different adaptive array configurations and higher L&S resolution.

In a document classification system, the collected omnidirectional or directional signals can be divided into two regions: the near-field (Fraunhofer region) and the radiative near-field (Fresnel region). The boundary between these two regions is usually defined by the plane-wave-based Fraunhofer distance (also known as the Rayleigh distance), which is given by

$$D = \frac{\lambda}{2\pi} \sqrt{r^2 + \left(\frac{\lambda}{2}\right)^2}$$

where D denotes the aperture of the antenna array, and λ is the signal wavelength. The main boundary between these regions can be defined differently depending on the application. For example, when the emitted signal propagates into the free space, the boundary is given by

$$d = \frac{\lambda}{2\pi} \sqrt{r^2 + \left(\frac{\lambda}{2}\right)^2}$$

which implies that the receiving antenna array can only resolve the source when it is located within the distance d .

A waveform model should be used to characterize the signal propagation within the near-field region, where the curvature of the wavefront is dominant. The wavefront curvature is used to determine both the direction and the propagation distance of the signal. Research on the L&S of sources or targets in the near-field region has been conducted for many years [15–19].

However, this research topic did not attract widespread attention until the late 1990s, when the interest in L&S increased.

At that time, the available signal processing techniques were still limited, rendering the sources or targets normally located in the far-field region.

(a') Οριζόντιο και κάθετο προφίλ

(β') Τελικό αποτέλεσμα

6 Κώδικας

```
1 import cv2
2 import numpy as np
3 from HW_1.Src.helper_functions import show_window_ratio, show_window
4 from HW_1.Src.integral import integral
5
6 DEBUG = False
7 # DEBUG = True # Just comment this out for ""release"""
8
9 # print(cv2.__version__)
10
11 def exercise(dut_array_image):
12     original_integral = integral(dut_array_image)
13
14     _, thresh = cv2.threshold(dut_array_image, 250, 255, cv2.THRESH_BINARY_INV)
15     kernel = cv2.getStructuringElement(cv2.MORPH_RECT, (31, 31))
16     morph_close = cv2.morphologyEx(thresh, cv2.MORPH_CLOSE, kernel)
17
18     show_window_ratio('morphology', morph_close.copy())
19
20     text_mask = np.zeros_like(thresh)
21     text_mask[thresh == 0] = 255
22
23     text_integral = integral(text_mask)
24     (totalLabels, label_ids, values, centroid) =
25         cv2.connectedComponentsWithStats(morph_close, 4)
26
27     detected = cv2.cvtColor(dut_array_image.copy(), cv2.COLOR_GRAY2BGR)
28
29     for i in range(1, totalLabels):
30         sub_area = values[i, cv2.CC_STAT_AREA]
31         x1 = values[i, cv2.CC_STAT_LEFT]
32         y1 = values[i, cv2.CC_STAT_TOP]
33         w1 = values[i, cv2.CC_STAT_WIDTH]
34         h1 = values[i, cv2.CC_STAT_HEIGHT]
35         bounding_box_area = w1 * h1
36
37         pt1 = (x1, y1)
38         pt2 = (x1 + w1, y1 + h1)
39
40         sub_image = detected[y1:(y1 + h1), x1:(x1 + w1)]
41         cv2.rectangle(detected, pt1, pt2, (0, 0, 255), 3)
42         cv2.putText(detected, f'Region {i}', (x1, y1 - 10), cv2.FONT_HERSHEY_SIMPLEX,
43             0.8, (0, 0, 255), 2)
44
45         sub_integral = original_integral[y1 + h1, x1 + w1] - original_integral[y1 - 1, x1] -
46             original_integral[
47                 y1, x1 - w1] + original_integral[y1 - 1, x1 - 1]
48         mean_grey = sub_integral / sub_area
```

```

46     sub_text_integral = text_integral[y1 + h1, x1 + w1] - text_integral[y1 - 1, x1] -
47         ↵ text_integral[y1, x1 - 1] + \
48             text_integral[y1 - 1, x1 - 1]
49
50
51 assert sub_area <= bounding_box_area
52
53 sub_thresh = thresh[y1:(y1 + h1), x1:(x1 + w1)]
54 # Κάνω morphological close για να ενώσω τα γράμματα σε λέξεις
55 word_kernel = cv2.getStructuringElement(cv2.MORPH_RECT, (5, 5))
56 word_close = cv2.morphologyEx(sub_thresh, cv2.MORPH_CLOSE, word_kernel)
57
58 # Μετράω τα connected components = λέξεις
59 word_count, _, sub_values, _ = cv2.connectedComponentsWithStats(word_close, 4)
60 # cv2.imwrite(f'./Report/text_segmentation/parts/{i}.png', word_close)
61
62 if DEBUG:
63     print(f'-----Region {i} -----')
64     print(f'Bounding box area(px): {bounding_box_area}')
65     print(f'Area (px): {sub_area}')
66     print(f'Words : {word_count}')
67     print(f'Mean gray-level value in bounding box: {mean_grey}')
68     # show_window(f'{i}.png', word_close)
69
70 show_window_ratio('detected', detected.copy())
71 # cv2.imwrite('./Report/text_segmentation/regions.png', detected)
72 return (detected,morph_close)

73 if __name__ == "__main__":
74     original_img = 'Images/5.png'
75     original = cv2.imread(original_img, cv2.IMREAD_GRAYSCALE)
76     exercise(original)

77 import cv2
78 from matplotlib import pyplot as plt
79 from HW_1.Src.filters import median_filter
80 from HW_1.Src.helper_functions import histogram_extract, show_window,
81     ↵ show_window_ratio
82 from noise_free import exercise

83 DEBUG = True

84
85 original_image_ctx = "Original Image"
86 original_image = "./Images/5_salt_pepper.png"
87 original = cv2.imread(original_image, cv2.IMREAD_GRAYSCALE)

88 assert original is not None

89 kernel_size = 5
90 filtered = median_filter(original, kernel_size)

91 final,morph = exercise(filtered)

```

```

19 assert final.any()
20 assert morph.any()
21 if DEBUG:
22     cv2.imwrite(f'./Report/salt_n_pepper/final_{kernel_size}.png', final)
23     cv2.imwrite(f'./Report/salt_n_pepper/morph_{kernel_size}.png', morph)

1 import cv2
2 from matplotlib import pyplot as plt
3 from HW_1.Src.filters import gaussian_filter
4 from HW_1.Src.helper_functions import histogram_extract, show_window,
   ↵ show_window_ratio
5 from noise_free import exercise

6
7 DEBUG = True

8
9 kernel_size = 31

10
11 original_image_ctx = "Original Image"
12 original_image = f"./Images/5.png"
13 original = cv2.imread(original_image, cv2.IMREAD_GRAYSCALE)
14 gaussian = gaussian_filter(original, kernel_size, 0)

15
16 assert original is not None

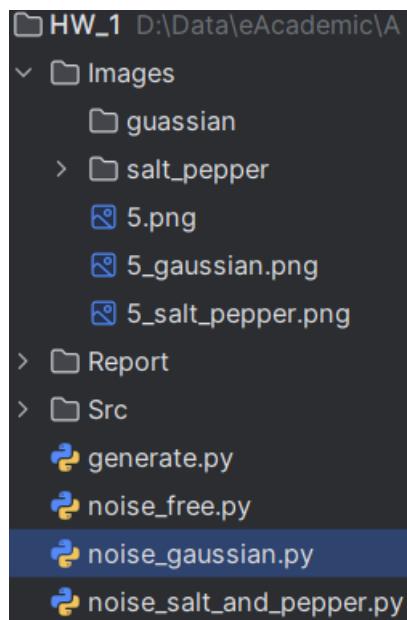
17
18 filtered = gaussian_filter(original, kernel_size)

19
20 final,morph = exercise(filtered)
21 assert final.any()
22 assert morph.any()
23 if DEBUG:
24     cv2.imwrite(f'./Report/gaussian/final_{kernel_size}.png', final)
25     cv2.imwrite(f'./Report/gaussian/morph_{kernel_size}.png', morph)

```

7 Δομή Project

7.1 Πως να τρέξετε τα αποτελέσματα



Το κάθε αρχείο μπορεί να τρέξει ανεξάρτητα, επιλέγοντας ένα και τρέχοντας το θα δείτε τα αποτελέσματα.

8 Αποτελέσματα

-----Region 1 -----

Bounding box area(px): 1684685

Area (px): 1332737

Words : 313

Mean gray-level value **in** bounding box: 326.8034668505489

-----Region 2 -----

Bounding box area(px): 2778

Area (px): 2778

Words : 2

Mean gray-level value **in** bounding box: 275788.83225341974

-----Region 3 -----

Bounding box area(px): 1541592

Area (px): 1260396

Words : 328

Mean gray-level value **in** bounding box: 720.7338899837829

-----Region 4 -----

Bounding box area(px): 383150

Area (px): 268532

Words : 74

Mean gray-level value **in** bounding box: 3872.8520399803374

-----Region 5 -----

Bounding box area(px): 13783

Area (px): 4322

Words : 7

Mean gray-level value **in** bounding box: 41702.24525682555

-----Region 6 -----

Bounding box area(px): 1428

Area (px): 1263

Words : 4

Mean gray-level value **in** bounding box: 38940.275534441804

-----Region 7 -----

Bounding box area(px): 810920

Area (px): 678866

Words : 163

Mean gray-level value **in** bounding box: 8.727071321880901

-----Region 8 -----

Bounding box area(px): 65190

Area (px): 21795

Words : 22

Mean gray-level value **in** bounding box: 32465.812525808673

-----Region 9 -----

Bounding box area(px): 1462

Area (px): 1263

Words : 3

Mean gray-level value **in** bounding box: 51022.54315122723

-----Region 10 -----

Bounding box area(px): 519384

Area (px): 402342

Words : 128

Mean gray-level value **in** bounding box: 3438.9615202986515

-----Region 11 -----

Bounding box area(px): 24654

Area (px): 15043

Words : 9

Mean gray-level value **in** bounding box: 57001.66828425181

-----Region 12 -----

Bounding box area(px): 870

Area (px): 810

Words : 3

Mean gray-level value **in** bounding box: 72722.96296296296

Region 1

manipulated to improve directionality and gain, which are crucial for L&S. The development of phased-array antennas for radar in the 1950s was a significant milestone [2]. In this era, a new initiative steered the direction of the beam electronically without moving the physical antennas, thereby greatly enhancing the speed and flexibility of L&S. With the advancement of antenna array technologies, array signal processing has gradually become an active research area since the 1980s [3]. This period, extending into the 1990s, witnessed a surge of classical L&S methods based on array signal processing, including the famous MUSIC (MUltiple SIgnal Classification), parametric maximum-likelihood estimator, and ESPRIT (Estimation of Signal Parameters via Rotational Invariance Technique) methods. These methods laid an important foundation for array signal processing techniques for L&S.

The turn of the century brought about a digital revolution, further transforming the capabilities of antenna arrays. The integration of digital signal processing with antenna array technology led to the development of adaptive antenna systems that are capable of dynamically adjusting their patterns to optimize L&S accuracy. This era also witnessed the advent of multiple-input, multiple-output (MIMO) systems [4], which use multiple antennas at both the transmitter and receiver to improve communication performance through beamforming, spatial multiplexing, and diversity. Building on these advancements, the concept of MIMO radar emerged in the early 2000s [5], allowing each antenna to transmit different waveforms and thus providing more adaptive array configuration and higher L&S resolution.

The electromagnetic (EM) radiation field emitted by antennas or antenna arrays can be divided into two regions: the far-field (Fraunhofer) region and the radiative near-field (Fresnel) region [6], [7]. The demarcation between these regions is traditionally defined by the phase error-based Fraunhofer distance (also known as the *Rayleigh distance*), which is given by

$$d_{FA} = \frac{2D^2}{\lambda}$$

Region 5

where D denotes the aperture of the antenna (array), and λ is the signal wavelength. The exact boundary between these regions can be defined differently depending on the application. Regardless, when the emitted signal propagates into the far-field region, its wavefront is approximated as planar, which implies that the receiving antenna array can only resolve the direction of the signal. However, the accurate spherical-wavefront model should be used to characterize the signal propagation within the near-field region, where the curvature is distance dependent. This property allows the antenna array to determine both the direction and the propagation distance of the signal. Research on the L&S of sources or targets in the near-field region dates back to as early as the 1980s [8], [9]. However, this research topic did not attract widespread attention during that era. This is because array apertures and signal frequencies were generally limited, rendering the sources or targets normally located in the far-field region.

Region 3

The landscape shifted with the introduction of the massive MIMO concept by Marzetta in 2010 [10]. The core idea of massive MIMO is to equip base stations with numerous antennas that are connected to fully digital transceiver branches, allowing them to generate many simultaneous narrow beams with high controllability. Massive MIMO was developed for communications but builds on enhancing the spatial resolution compared to smaller arrays, which also significantly improves L&S accuracy. These favorable features have made an elementary form of massive MIMO with 32–64 antennas the fundamental building block of 5G networks and typically operate in the 3.5-GHz band. Since the emergence of 5G massive MIMO, researchers have continued to envision using even larger arrays with various form factors and characteristics, inevitably increasing the physical size of the antenna arrays. Coupled with the trend toward exploiting high-frequency bands (centimeter and millimeter waves and beyond), which have very short wavelengths, the near-field region around base stations can be significantly expanded. For instance, an antenna array with a 1-m aperture operating at 30 GHz can have a near-field region extending up to 200 m, making it highly likely for targets to be located within this near-field region. Against the aforementioned background, a lot of research efforts have been devoted to the field of near-field L&S, focusing on various aspects such as signal modeling [6], [7], [11], [12], signal processing [13], [14], [15], [16], [17], [18], [19], and performance-bound characterization [20], [21]. This article aims to provide a tutorial review on near-field L&S from signal modeling to signal processing. Our goal is to highlight the most significant changes from traditional far-field L&S and elucidate the key intuitive ideas for leveraging the advantages of near-field L&S and addressing the associated challenges.

Region 4**Near-field properties**

In this section, we first outline near-field propagation from the perspective of EM theory and discuss its relevance to the near-field L&S considered in this article. Let us consider a sinusoidal wave with frequency f_c transmitted from a point source. The Fourier representation of the electric field observed at a distance r from the source, in any direction perpendicular to the propagation direction, is proportional to [12]

$$G(r) = -\frac{j\eta e^{-j\frac{2\pi}{\lambda}r}}{2\lambda r} \left(1 + \frac{j\lambda}{2\pi r} - \frac{\lambda^2}{(2\pi r)^2} \right) \quad \boxed{2}$$

Region 8

where η denotes the intrinsic impedance, $\lambda = c/f_c$ denotes the wavelength with c being the speed of light, and $j = \sqrt{-1}$. The last two terms, which decay faster than $1/r$, diminish rapidly as r increases and do not contribute to EM radiation. These terms are associated with “near-field” propagation in EM theory, but they are more commonly referred to as *reactive near-field* in recent literature on near-field communications and L&S. It has been demonstrated that the reactive near field is confined within a few wavelengths from the source, even for an infinitely large antenna (or array) [22]. Therefore, in practical system design, only the radiating field

$$\boxed{6}$$

$$\boxed{1}$$

Region 9

$$\boxed{2}$$

Region 10

where η denotes the intrinsic impedance, $\lambda = c/f_c$ denotes the wavelength with c being the speed of light, and $j = \sqrt{-1}$. The last two terms, which decay faster than $1/r$, diminish rapidly as r increases and do not contribute to EM radiation. These terms are associated with “near-field” propagation in EM theory, but they are more commonly referred to as *reactive near-field* in recent literature on near-field communications and L&S. It has been demonstrated that the reactive near field is confined within a few wavelengths from the source, even for an infinitely large antenna (or array) [22]. Therefore, in practical system design, only the radiating field

Region 11**Region 12**

Αναφορές

- [1] NumPy Developers. (2024). *numpy.cumsum* (NumPy Stable Manual). Retrieved October 28, 2024, from <https://numpy.org/doc/stable/reference/generated/numpy.cumsum.html>
- [2] Federico Boletti, Stefano Allegretti, and Costantino Grana. One dag to rule them all. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021.
- [3] Gonzalez, R. C. & Woods, R. E. *Digital Image Processing*. Prentice Hall, Κεφάλαιο 3.6: “Order-Statistic (Nonlinear) Filters”.
- [4] OpenCV Official Documentation, Structural Analysis and Shape Descriptors https://docs.opencv.org/4.11.0/d3/dc0/group__imgproc__shape.html#ga5ed7784614678adccb699c70fb841075
- [5] OpenCV Official Documentation, Otsu’s Binarization https://docs.opencv.org/4.x/d7/d4d/tutorial_py_thresholding.html
- [6] OpenCV Official Documentation, Thresholding https://docs.opencv.org/4.x/d7/d4d/tutorial_py_thresholding.html