

# **APLIKASI PENGENALAN HURUF ALFABET BAGI ANAK USIA DINI MENGUNAKAN METODE PENGOLAHAN CITRA BERBASIS DATA SUARA**

---

---

**Betty Dewi Puspasari<sup>14</sup>**

## **Abstrak**

Pengenalan suara merupakan salah satu media pembelajaran yang dapat mendukung proses belajar mengajar, terutama proses pengenalan huruf alfabet bagi anak usia dini.

Pada penelitian ini dilakukan dengan langkah-langkah berupa mendapatkan data suara dari ahli pembelajaran dan suara anak yang diubah dalam bentuk grafik. Dari data grafik yang berbentuk citra gambar tersebut dilakukan normalisasi intensitas. Dari nilai normalisasi intensitas dilakukan pencocokan pola dengan metode *pixel matching*.

Aplikasi pengenalan huruf alfabet ini menampilkan data dalam audio dan visual dalam bentuk histogram dari grafik. Dengan menggunakan tampilan data dalam Audio visual dan gambar grafik, dapat membandingkan pengucapan yang benar dan anak usia dini. Perbandingan dapat dilakukan dengan dua cara. Perbandingan pertama adalah perbandingan audio visual, secara keseluruhan hasil yang diperoleh 73,8% benar dan 26,2% tidak benar. Perbandingan yang kedua dengan menggunakan metode *Image Processing*, hasil keseluruhan yang diperoleh, dengan menggunakan kesamaan itu 74,5% benar dan 25,5% tidak benar.

**Kata-kata kunci:** alfabet, pengolahan citra

## **Abstract**

*Speech recognition is one of the media that can support teaching and learning, particularly the introduction of the alphabet for children.*

*This research performs some steps such as getting data from the noise and the voice of children learning expert who converted into the form of graphs. The data in the form of graphic image undergo intensity*

---

<sup>14</sup> Betty Dewi Puspitasari. Dosen Jurusan Teknik Informatika, Sekolah Tinggi Teknik Atlas Nusantara Malang.

*normalization. From normalized intensity values pattern matching method is done using pixel matching method.*

*This alphabet recognition application displays audio and visual data in the form of a histogram graph. Using the display of data in audio visual and graphic images, to compare the correct pronunciation and early childhood. Comparison can be done in two ways. The first comparison is audio visual comparison, the overall results obtained 73.8% and 26.2% completely untrue. The second comparison using the image processing, the overall results obtained, using the similarity was 74.5% and 25.5% completely untrue.*

**Keywords:** *Alphabet, Image Processing*

## **1. PENDAHULUAN**

Pola kurikulum sebelum dilakukan perubahan kurikulum pendidikan dasar 2013 di Indonesia kurang baik bagi anak-anak awal sekolah dasar, pada kelas awal terutama kelas 1 dan 2 SD merupakan dasar untuk memperoleh kemampuan bahasa secara baik. Namun, pendidik menjadi salah kaprah tidak lagi memperhatikan kemampuan anak usia dini yang baru mengenal baca dan tulis di masa prasekolah. Guru kelas 1 SD menginginkan agar anak didiknya harus sudah bisa baca dan tulis dengan baik. Guru kelas 1 SD menganggap guru TK yang harus membelajarkan baca dan tulis. Sementara dalam kurikulum TK kegiatan baca dan tulis bukan merupakan fokus bagi pembelajaran di TK. Mereka perlu diperkenalkan huruf, angka dan membaca secara umum bukan menjadi titik fokus yang harus dikuasai oleh anak TK.

Kemajuan teknologi bertujuan untuk mempermudah kegiatan manusia di segala bidang. Pemanfaatan teknologi informasi telah merambah semua bidang dan tak terkecuali bidang pendidikan. Dengan adanya kemajuan teknologi, pembuatan media pembelajaran dapat dibuat menjadi lebih menarik. Melalui pembelajaran bahasa yang menyenangkan diharapkan dapat membantu anak untuk memperoleh kemampuan melatih penerapan pengucapan huruf alfabet.

Pengenalan suara merupakan salah satu bagian dari bidang aplikasi yang memungkinkan suatu perangkat untuk mengenali dan memahami kata kata yang diucapkan dengan cara digitalisasi kata dan pencocokan sinyal digital tersebut dengan suatu pola

tertentu yang tersimpan dalam suatu perangkat. Kata-kata yang diucapkan diubah bentuknya menjadi sinyal digital dengan cara mengubah gelombang suara menjadi sekumpulan angka yang kemudian disesuaikan dengan kode-kode tertentu untuk mengidentifikasi kata-kata tersebut. Hasil dari identifikasi kata yang diucapkan dapat ditampilkan dalam bentuk tulisan atau dapat dibaca oleh perangkat teknologi.

Hasil pengolahan sinyal menghasilkan angka-angka yang dikumpulkan dalam sebuah grafik. Dari grafik yang terbentuk dilakukan pencocokan pola sehingga akan didapat apakah pengucapan yang dilakukan oleh suara anak sudah benar apa belum. Pencocokan pola menggunakan metode pada pengolahan citra yaitu metode preprocessing citra berupa normalisasi citra yang didapatkan dari data citra suara pembelajaran dan data citra suara anak serta metode *pixel matching* antara kedua data citra tersebut untuk menghasilkan kesimpulan kebenaran pengucapan benar atau salah.

## **2. KAJIAN PUSTAKA**

### **2.1 Data Suara**

Data Suara dalam bentuk File WAV adalah file audio standar yang digunakan oleh Windows. Suara yang berupa digital audio dalam file WAV disimpan dalam bentuk gelombang, karena itulah file ini memiliki ekstensi .WAV. File WAV ini dapat dibuat dengan menggunakan berbagai wave editor maupun wave recorder, misalkan Sound Recorder milik Windows.

Data digital audio dalam file WAV bisa memiliki kualitas yang bermacam-macam. Kualitas dari suara yang dihasilkan ditentukan dari bitrate, sample rate, dan jumlah channel.

Bitrate merupakan ukuran bit tiap sampelnya, yaitu 8-bits, 16-bits, 24-bit atau 32-bits. Dalam 8-bits WAV semua sample hanya akan memakan sebanyak 1 byte saja. Sedangkan untuk 16-bits tiap sampelnya memiliki nilai antara -32768 sampai 32767, dibandingkan untuk 8-bits yang hanya memiliki nilai antara -128 sampai 127 maka 16-bits WAV file menghasilkan suara yang lebih baik karena datanya lebih akurat daripada 8-bits.

Samplerate menyatakan banyaknya jumlah sample yang dimainkan setiap detiknya. Samplerate yang umum dipakai

adalah 8000Hz, 11025Hz, 22050Hz, dan 44100Hz. Untuk 8000Hz biasanya disebut sebagai Telephone quality karena suara yang dihasilkan menyerupai suara dari telepon. Samplerate 11025Hz sering digunakan untuk merekam suara manusia karena dinilai sudah cukup. Tetapi untuk musik agar dapat memperoleh hasil yang baik maka digunakan samplerate 22050Hz. Sedangkan 44100Hz merupakan CD quality, samplerate 44100Hz inilah yang digunakan dalam audio CD karena bagus untuk semua jenis suara.

Berbeda dengan bitrate, perbedaan samplerate dapat dikenali dengan mudah, karena tentu saja suara dari radio dengan suara CD player sangat jauh berbeda.

Jumlah channel menentukan suara yang dihasilkan apakah mono atau stereo. Mono memiliki hanya 1 channel, sedangkan stereo 2 channel dan memakan tempat 2 kali lebih banyak daripada mono. Untuk merekam suara manusia, mono sudah cukup memberikan kualitas yang baik.

### **Struktur File WAV**

File WAV menggunakan struktur standar RIFF dengan mengelompokkan isi file ke dalam format WAV dan data digital audio. Setiap bagian memiliki header sendiri beserta ukurannya. Struktur RIFF (Resource Interchange File Format) adalah sebuah struktur yang umumnya digunakan untuk data multimedia di Windows. Struktur ini mengatur data dalam file ke dalam bagian-bagian yang masing-masing memiliki header dan ukurannya sendiri dan disebut chunk. Struktur ini memungkinkan bagi program bila tidak mengenali bagian tertentu untuk melompati bagian tersebut dan terus memproses bagian yang dikenal. Data suatu bagian bisa memiliki sub-bagian dan seluruh data dalam file berstruktur RIFF selalu merupakan sub-bagian dari suatu bagian yang memiliki header "RIFF". Contoh file yang menggunakan struktur RIFF adalah file WAV dan AVI.

## **2.2 Pengolahan Citra**

Langkah dalam proses pengolahan citra secara umum terdiri dari bagian seperti:

- a. Pembentukan Citra (*Data Acquisition*), menentukan data yang diperlukan dan memilih metode perekaman citra *digital*.
- b. Pengolahan Citra Tingkat Awal (*Image Preprocessing*), meningkatkan kontras, menghilangkan gangguan geometrik/radiometrik, menentukan bagian citra yang akan diobservasi. Segmentasi Citra (*Image Segmentation*) dan Deteksi Sisi (*Edge Detection*): Melakukan partisi citra menjadi wilayah-wilayah obyek (*internal properties*) atau menentukan garis batas wilayah obyek (*external shape characteristics*).
- c. Seleksi dan Ekstraksi Ciri (*Feature Extraction and Selection*), seleksi ciri memilih informasi kuantitatif dari ciri yang ada, yang dapat membedakan kelas-kelas obyek secara baik. Ekstraksi ciri mengukur besaran kuantitatif ciri setiap *Pixel*.

### **Segmentasi Berdasarkan Fitur Warna**

Segmentasi warna bekerja dengan mengenali informasi warna dari setiap *Pixel* dan mengelompokkannya sesuai fitur yang diinginkan dan tingkat kesamaannya. Pendekatan segmentasi yang umum dipakai adalah berbasis intensitas, berbasis warna, dan berbasis bentuk.

### **Model YCbCr**

Bentuk YCbCr banyak digunakan dalam *video digital*. Dalam format ini, informasi *luminance* disimpan dalam komponen tunggal (Y), dan informasi *chrominance* disimpan dalam dua komponen warna (Cb dan Cr). Cb merepresentasikan perbedaan antara komponen warna biru dan nilai referensi, dan Cr merepresentasikan perbedaan antara komponen warna merah dan nilai referensi.

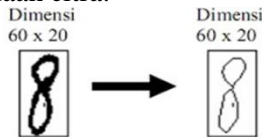
Informasi *luminance* disimpan dalam komponen tunggal (Y), dan informasi *chrominance* disimpan dalam dua komponen warna (Cb dan Cr). Cb merepresentasikan perbedaan antara komponen warna biru dan nilai referensi, dan Cr merepresentasikan perbedaan antara komponen warna merah dan nilai referensi.

Persamaan RGB ke dalam YCbCr adalah seperti pada rumus (1).

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} + \begin{bmatrix} 65.481 & 128.553 & 24.966 \\ -37.797 & -74.203 & 112.000 \\ 112.000 & -93.786 & -18.214 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (1)$$

### Preprocessing-Normalisasi

Citra dipetakan pada pixel dengan ukuran tertentu sehingga memberikan representasi dimensi yang tetap. Tujuan dari normalisasi citra adalah mengurangi resolusi citra yang berguna saat proses pengenalan citra dan juga meningkatkan akurasi pengenalan. Proses yang digunakan pada tahap normalisasi ini adalah proses penskalaan citra.



Gambar 1. Visualisasi Normalisasi Ketebalan

Untuk menyeragamkan pola-pola karakter yang akan digunakan sebagai input jaringan syaraf tiruan, maka ketebalan dari tiap-tiap pola diseragamkan. Dalam penelitian ini, normalisasi ketebalan ini diseragamkan seolah-olah mendekati 1 piksel. Gambar citra yang mengalami normalisasi ketebalan tampak seperti pada gambar 1.

### 2.3 Pengenalan Pola

Pengenalan pola dapat dikatakan sebagai kemampuan mengenali citra berdasarkan ciri-ciri dan pengetahuan yang pernah diamatinya dari citra-citra tersebut. Tujuan dari pengenalan pola adalah mengklasifikasi dan mendeskripsikan pola atau citra kompleks melalui pengetahuan sifat-sifat atau ciri-ciri citra tersebut.

Ada tiga pendekatan dalam pengenalan pola yaitu secara sintaks, statistik, dan semantik[5]. Pengenalan pola secara sintaks dilakukan berdasarkan ciri-ciri objek. Pengenalan pola secara statistik dilakukan berdasarkan komputasi matematis. Pendekatan dengan semantik berarti pola dikenali dalam tataran yang lebih abstrak.

Tabel 1.Sistem Aturan Bahasa

No	Sistem Aturan	Kejadian
1	Fonologi	Kesulitan untuk mengucapkan kelompok konsonan misalnya strika. Umumnya akan berlanjut hingga memasuki masa akhir anak-anak.
2	Morfologi	Anak sudah dapat mengungkapkan lebih dari dua kata misalnya kakak memukul saya dan saya dipukul kakak.
3	Sintaksis	Anak belajar menerapkan aturan untuk membentuk kalimat sesuai aturan. Misalnya badu membawa buku bukan membawa badu buku. Orang tua memperbaiki kalimat yang salah tersebut untuk sesuai aturan SPOK. Akan terus berlanjut pembelajaran hingga masa akhir dengan kalimat yang lebih kompleks.
4	Semantik	Anak pada masa ini sangat pesat belajar kosa kata hingga 5 - 8 kata sehari mereka serap sehingga diperkirakan pada usia 6 tahun sudah 8000 hingga 14000 kata yang sudah dikuasai.
5	Pragmatik	Usia 2 tahun berbeda dengan 6 tahun, pada usia 6 tahun mereka sudah dapat berbicara imajinatif dan gaya bicara lebih sopan dengan orang dewasa.

Dikutip dari Ralph Waldo dalam Santrock (2007:355)

## 2.4 Pembelajaran Bahasa Anak Usia Dini

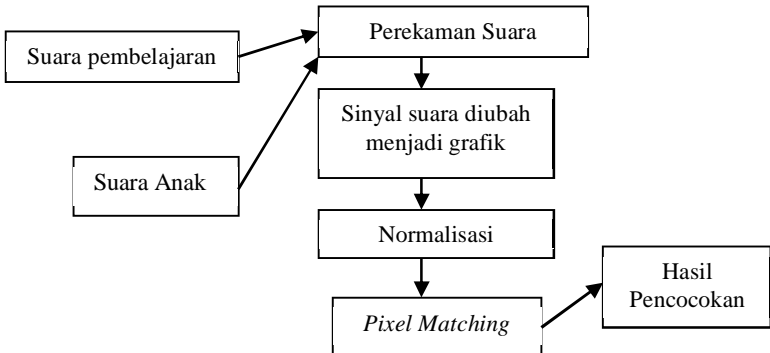
Manusia sebagai makhluk sosial tidak bisa dipisahkan dari kegiatan saling berkomunikasi. Untuk berkomunikasi manusia memerlukan suatu media yaitu bahasa. Bahasa adalah suatu bentuk komunikasi lisan, tertulis atau isyarat yang digunakan secara kombinasi oleh masyarakat. Bahasa merupakan suatu kode atau sistem simbol dan urutan kata-kata yang diterima secara konvensional untuk menyampaikan konsep- konsep atau ide-ide dan berkomunikasi melalui penggunaan simbol-simbol yang diatur oleh ketentuan yang ada. Bahasa juga merupakan alat komunikasi dengan orang lain dan kemudian berlangsung dalam suatu interaksi sosial. Bahasa meliputi empat aspek keterampilan yaitu aspek keterampilan berbicara, menulis, menyimak, dan membaca. Ada lima sistem aturan bahasa dapat dilihat pada Tabel 1.

### 3. METODE

#### 3.1 Teknik Analisis Data

Teknik analisa data disini didasarkan pada metode *speech recognition* yaitu pembacaan sinyal suara oleh komputer yang direpresentasikan dalam bentuk tampilan grafik sinyal suara dan tampilan suara. Grafik dan suara yang dihasilkan oleh suara dalam library dan suara anak akan dianalisa berdasarkan kemiripan atau similaritas secara audio dan visual dan secara pengolahan citra (*image processing*).

#### 3.2. Blok Diagram



Gambar 2. Blok Diagram Penelitian

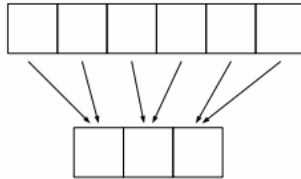
Pada blok Penelitian (Gambar 2) dapat dilihat bahwa pengolahan sinyal suara dari suara pembelajaran dan anak yang masuk dalam aplikasi komputer. Data Suara yang diambil adalah data suara pembelajaran dan data suara anak, yang selanjutnya mengalami proses perekaman. Hasil perekaman tersebut berupa file yang berekstensi .wav yang merupakan data yang akan ditampilkan pada aplikasi *speech recognition*. Pada aplikasi *speech recognition* file data yang sudah masuk ditampilkan dalam bentuk grafik (visual) dan audio. Tahap selanjutnya adalah tahap proses dimana pada tahap ini terdapat dua cara membandingkan



yaitu dengan cara audio dan visual dan dengan cara *image processing*. Tahap pencocokan adalah tahap akan mencocokkan grafik suara pembelajaran dengan grafik suara anak. Hasil pencocokan adalah merupakan hasil kemiripan grafik suara anak dan grafik suara pembelajaran akan muncul nilai similaritas dari dua citra yang dibandingkan.

### 3.3 Algoritma Penskalaan

Scaling atau penskalaan pada citra disebut juga *image zooming*, yaitu proses untuk mengubah ukuran citra asli (*zoom in* / memperbesar ukuran citra asli atau *zoom out* / memperkecil ukuran citra asli). Proses perubahan ukuran resolusi citra dibutuhkan untuk menyesuaikan resolusi citra masukan dengan resolusi citra template.



Gambar 3. Perubahan ukuran pixel.

Gambar 3 mengilustrasikan perubahan ukuran pixel dengan ukuran awal 6x1 pixel menjadi 3x1 pixel. Garis panah yang berasal dari pixel lama ke pixel baru mengilustrasikan pemetaan koordinat setiap pixel lama terhadap pixel baru. Persamaan 2 digunakan untuk menentukan faktor pengali dalam menentukan koordinat y pada pixel baru.

$$dy = \frac{\text{newheight}}{\text{tinggi} + 1} \quad (2)$$

Persamaan 3 digunakan untuk menghitung faktor pengali dalam menentukan koordinat x.

$$dx = \frac{\text{newwidth}}{\text{lebar} + 1} \quad (3)$$

Persamaan 4 digunakan untuk menentukan koordinat x.

$$x = i * dx \quad (4)$$

Persamaan 5 digunakan untuk menentukan koordinat  $y$ .

$$y = i * dy \quad (5)$$

### 3.4 Pixel Matching

Pada dasarnya pixel matching adalah proses yang sederhana. Suatu citra masukan tertentu dibandingkan dengan citra pada basis data. Pusat bagian citra yang akan dibandingkan dan dihitung seberapa banyak titik yang paling sesuai dengan citra. Langkah ini diulangi terhadap keseluruhan citra masukan yang akan dibandingkan. Nilai kesesuaian titik yang paling besar antara citra masukan dan data citra menandakan bahwa data citra tersebut merupakan data citra yang paling sesuai dengan citra masukan.

Tingkat kesesuaian antara citra masukan dan data citra bisa dihitung berdasarkan nilai eror terkecil [6] dengan menggunakan persamaan 6.

$$\min e = \sum_{(x,y) \in W} (I_{x,y} - T_{x,y})^2 \quad (6)$$

$I$  adalah pola pixel citra masukan yang akan dibandingkan.  $T$  adalah pola pixel citra template. Template dengan nilai eror paling kecil adalah template yang paling sesuai dengan citra masukan yang akan dibandingkan. Ukuran objek yang beragam bisa diatasi dengan menggunakan template berbagai ukuran. Namun hal ini membutuhkan tambahan ruang penyimpanan.

Penambahan data citra dengan berbagai ukuran akan membutuhkan komputasi yang besar. Jika suatu template berukuran persegi dengan ukuran  $m \times m$  dan 19 sesuai dengan citra yang berukuran  $N \times N$ , dan dimisalkan pixel  $m^2$  sesuai dengan semua titik citra, maka komputasi yang harus dilakukan adalah  $O(N^2 m^2)$ .

Komputasi tersebut harus dilakukan dengan template yang tidak beragam. Jika parameter template bertambah, seperti ukuran template yang beragam, maka komputasi yang dilakukan juga akan bertambah. Hal ini yang menyebabkan metode template matching menjadi lambat.

#### 4. HASIL DAN PEMBAHASAN

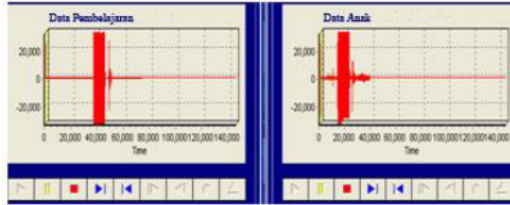
Ada dua tampilan data dalam penelitian ini, yang pertama adalah suara pembelajaran dan suara anak secara audio, yang ditampilkan menggunakan media player, dan yang kedua adalah grafik dari masing masing suara tersebut.

Aplikasi ini sudah memiliki database data citra untuk semua jenis alpabet yang didapatkan dari ahli materi. Pada saat pertama aplikasi ini dijalankan akan ditampilkan terlebih dahulu salah satu jenis alfabet yang dikeluarkan secara acak yang ditampilkan dalam bentuk obyek Teks dalam aplikasi. Langkah selanjutnya pengguna dalam hal ini seorang anak diminta untuk mengucapkan huruf yang terdapat dilayar. Setelah anak tersebut memasukkan suara lewat *microphone*, kemudian hasil suara anak yang tersimpan dalam format WAV sehingga dapat didapatkan data sinyal suara. Sinyal suara tersebut akan dikonvert menjadi sinyal suara dalam bentuk data citra gambar.

Data citra gambar yang didapatkan selanjutnya dilakukan normalisasi dalam ukuran 150x150 pixel. Langkah selanjutnya dilakukan pencocokan antara data pembelajaran dengan data suara anak dengan metode pixel matching. Pada akhirnya akan didapatkan tingkat kecocokan antara data pembelajaran dengan data citra anak. Untuk menentukan pengucapan anak tersebut benar atau salah.



Gambar 4. Tampilan Aplikasi



Gambar 5. Contoh proses Membandingkan Grafik Secara Audio Visual oleh Salah Satu Anak pada Huruf “D”

Proses ujicoba yang dilakukan adalah ujicoba lapangan yaitu dilaksanakan pada 30 anak di paud wonderland malang. Ujicoba ini dilakukan dimana masing-masing anak diminta untuk mengucapkan huruf alfabet mulai dari 1 huruf sampai dengan variasi 5 huruf, total ujicoba permasing-masing anak sebanyak 10 ujicoba. Berdasarkan ujicoba yang dilakukan terhadap 30 anak. Didapatkan hasil seperti pada Tabel 2.

Tabel 2. Hasil Perbandingan

<b>Audio Visual</b>	<b>Jumlah</b>	<b>Persen</b>
Total benar	1107	73,8
Total tidak benar	393	26,2
<b>Image Processing</b>	<b>Jumlah</b>	<b>Persen</b>
Total benar	1117	74,5
Total tidak benar	383	25,5

## 5. PENUTUP

Aplikasi pengenalan huruf alfabet ini adalah salah satu aplikasi alternatif dalam membantu anak untuk mempelajari huruf alfabet. Dalam uji coba yang telah dilakukan didapatkan hasil tingkat keberhasilan aplikasi ini sebesar 74,5% dengan ujicoba pada 30 anak di Kota Malang.

## 6. DAFTAR PUSTAKA

Depdiknas, 1997. Metodik khusus pengembangan kemampuan bahasa di TK. Jakarta: Proyek Peningkatan Mutu TK Jakarta

- Depdiknas, 1997. Metodik umum di TK. Jakarta: Proyek Peningkatan Mutu TK Jakarta Depdiknas. (2004). Kurikulum berbasis kompetensi untuk jenjang TK. Jakarta: Pusat Kurikulum.
- Depdiknas, 2006. Kurikulum Tingkat Satuan Pelajaran untuk jenjang SD. Jakarta: Pusat Kurikulum.
- Dwijawiyata, 1970. Cakap membaca dan menulis. Yogyakarta: Kanisius
- Gonzales, R.C., and Woods, R.E., 2002. Digital Image Processing Second edition. New Jersey: Prentice Hall.
- Acharya, Tinku., and Ray, Ajoy K. 2005. Image Processing: Principles and Applications. New Jersey: John Wiley & Sons.
- Waniak, W. Image restoration by simple adaptive deconvolution. *Astron. Astrophys. Suppl. Ser.* 124, 197- 203
- Pantin, Eric. Starck, Jean-Luc. and Murtagh, Fionn. Deconvolution and Blind Deconvolution in Astronomy. CRC Press
- Caron, James N. 2002. Non-iterative blind data restoration using an extracted filter function. *OCIS.* 100.1830, 100.3020, 100.2000