

Classical versus Bayesian interferometric imaging of exoplanets

Tomasz Sołtysiński^{1,2†}

¹Nicolaus Copernicus Astronomical Centre, Polish Academy of Science
ul. Bartycka 18, 00-716 Warsaw, Poland
email: tomasz.s@camk.edu.pl

²Department of Mechatronics, Warsaw University of Technology
ul. Sw. Andrzeja Boboli 8, 02-525, Warsaw, Poland

Abstract. The interferometry is the most promising way to directly observe exoplanets, their spectra and surfaces at optical or infrared domain. The complex imaging process can be described as the extraction of information from the data gathered by the interferometer. This information can be treated to be independent on any *à priori* knowledge or the integration process. In this case it is analyzed in a classical way. In fact, the imaging of exoplanets is not the classical way of the data reconstruction. The best extraction of information requires all accessible *à priori* knowledge. This is the bayesian way. The knowledge gathered during an integration is also contributing as *à priori* information for futher image reconstruction. We discuss both approaches supporting the analysis with the estimates of information flow through the interferometer and examples of simulations of imaging.

Keywords. instrumentation: interferometers, methods: data analysis, techniques: image processing, astrobiology

1. Introduction

The motivation for this investigation comes from multidisciplinary point of view on exoplanet imaging science. Interferometers seem to be the best tools to image exoplanet (Angel & Woolf (1986), Bracewell (1978)). Serving first as a detectors of exoplanets in IR and visible domain (Léger (1999)), they are expected to provide first resolved images of exoplanets surfaces and their features (Labeyrie (1995)). This will allow the detailed investigation of their properties leading to the determination of life signs (Woolf & Angel (1998), Des Marais *et al.* (2002)). The main reason for this work is to understand the way how the interferometer process the information and hence, how to recover this information, classically or in a Bayesian way. We simulate the interferometric imaging process and investigate the information flow through different steps.

2. Simulations of exoearths interferometric imaging

2.1. Observing conditions

We start the simulation with the parameters typical for Earth-like planet placed at 10 pc from the Solar System. The interferometr is pointed towards the planet which is 1 AU far away from its host star. This results in planet angular size of $2.0645 \cdot 10^{-11}$ rad or $8.5 \mu\text{arcsec}$. The field of view (FoV) of interferometer was set to $40 \mu\text{arcsec}$ yielding the ratio $\text{FoV}/D \approx 4.7$, where D is planet angular diameter. The star flux is completely

† Present address: Dept. of Mechatronics, Warsaw University of Technology, Św. Andrzeja Boboli 8, 02-525 Warszawa

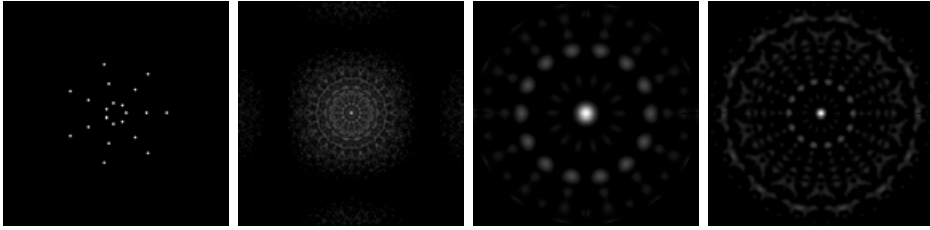


Figure 1. From left to right: the interferometer array design, its Fourier transform, the corresponding optical transmission function or interferometer response function for R and B band.

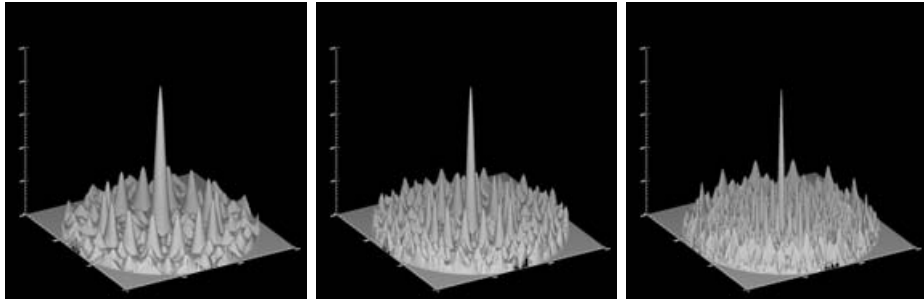


Figure 2. 3D shape of interferometer response function for R,G,B bands.

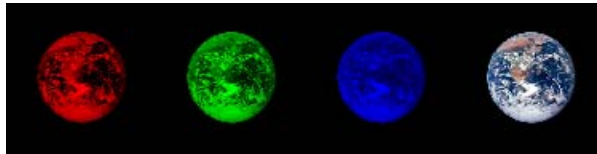


Figure 3. The flux pattern. The multiwavelength image of Earth taken by Apollo crew.

removed and has no influence on the reconstructed image. To simulate the flux coming from the distant planet we adapted the image of Earth as taken from space by Apollo crew (Fig. 3). The natural RGB components are then separated and serving as an image of exoearth for particular bands.

2.2. Interferometer setup

We have designed a space interferometer composed of sparse array of free flying telescopes. The array forms a pattern that corresponds to seven arms star with 3 rings (left of Fig. 1). The most distant spaceships are placed at 75 km from the centre, yielding 150 km of synthesized aperture. Such design results in interferometer response function or square modulus of Fourier transform of the initial array as shown on Fig. 1, right. Such optical transmission function yields

the patterns given on Fig. 2, with left picture for R (800 nm) band and next one for B (400 nm). The corresponding interferometer response functions (Absil (2001)) are shown as 3D plots on Fig. 2 for all 3 bands. The interferometer is not allowed to rotate, since we are interested only in its general information processing and not in the detailed (u, v) coverage necessary for accurate imaging. The instrument also stay tuned for the initial baseline length for all bands. No delay lines are assumed nor phase shifts introduced or we assume that all occurring phase shifts are compensated by achromatic phase shifters.

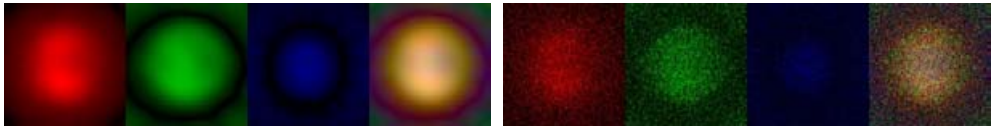


Figure 4. The data provided by the instrument for R,G,B bands and combined multichannel data on the left. The same data affected by instrumental noise with SNR=10.

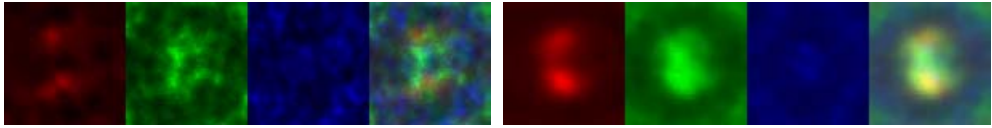


Figure 5. The MEM (left) and ML (right) deconvolved image of noise affected data after 50 iterations.

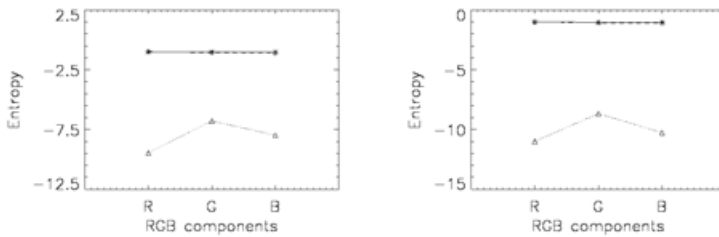


Figure 6. Multiscale entropy $H(X)$ for noise free data reconstruction compared with the multiscale entropy of initial data. Solid line is the original Apollo image, dashed – convolved interferometric data, dotted the deconvolved one, MEM, left and ML, right. All data is normalized in the way that absolute value of Apollo data’s maximum yields always 1.

2.3. Image content processing and extraction of information

2.3.1. Standard entropy

The information content of any signal or image is traditionally measured by its entropy, a term coming from

thermodynamics and communication theory Shanon (1948). It is expressed as

$$H_S(X) = \sum_{k=1}^N p_k \log(p_k) \tag{2.1}$$

where p_k is the probability of occurrence of pixel k among all N pixels of image X and typically taken from histogram.

The information obtained by investigation of entropy may have different interpretation and usually depend on the image content. In astronomy, the flat image has usually very low Shannon entropy and very noisy image has high entropy. We usually look for the level of entropy that combines the high image information content.

or desired information, with the lowest possible level of noise and disturbance coming from point spread function (PSF) of the instrument that served to provide the image. Having the entropy as common measure of image ordering we use it as a tool to estimate the information flow through the interferometer and reconstruction process. To have another measure of this process we also adapt Burg entropy that has slightly different

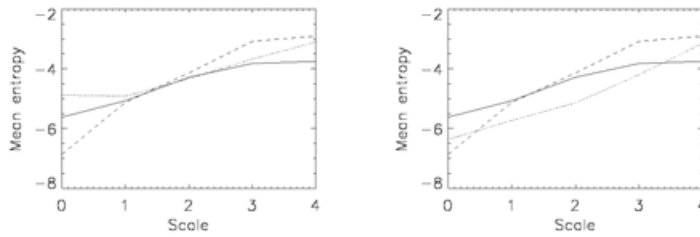


Figure 7. Mean entropy E_j vector for R component for all scales in the case of noise free simulation. Solid line is the original Apollo image, dashed – convolved data, dotted the deconvolved one, MEM, left and ML, right.

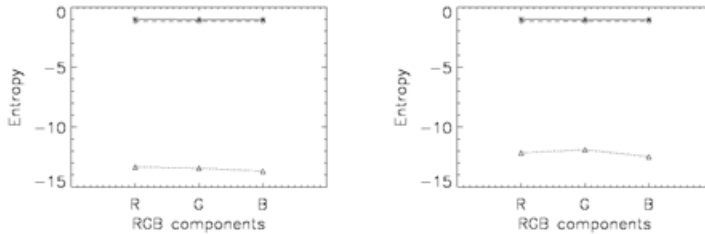


Figure 8. Multiscale entropy $H(X)$ for noise affected data reconstruction compared with the multiscale entropy of initial data. Solid line is the original Apollo image, dashed – convolved data, dotted the deconvolved one, MEM, left and ML, right. All data is normalized in the way that absolute value of Apollo data’s maximum yields always 1.

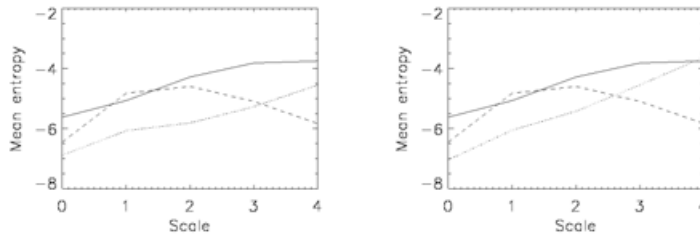


Figure 9. Mean entropy E_j vector for R component for all scales in the case of noise affected simulation. Solid line is the original Apollo image, dashed – convolved data, dotted the deconvolved one, MEM, left and ML, right.

definition

$$H_B(X) = - \sum_k \sum_l \ln(X(k,l)). \tag{2.2}$$

Another definitions of entropy are also widely used but we found them less useful for the purpose of our study.

2.3.2. *Multiscale entropy*

Since the introduction of wavelets into contemporary science a powerful tool of signal and image analysis has become accessible. The power of this approach comes from the orthogonality of the wavelet base functions and extremally wide range of their diversity,

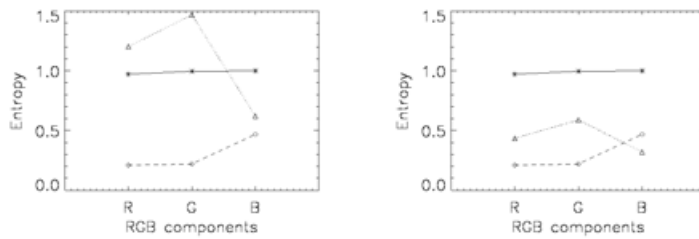


Figure 10. Normalized standard Shannon entropy of noise free case. Solid line is the original Apollo image, dashed – convolved data, dotted the deconvolved one, MEM, left and ML, right. All data is normalized in the way that absolute value of Apollo data’s maximum yields always 1.

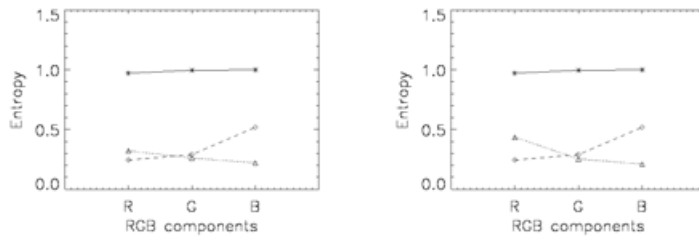


Figure 11. Normalized standard Shannon entropy H_S of noise affected case. Solid line is the original Apollo image, dashed – convolved data, dotted the deconvolved one, MEM, left and ML, right. All data is normalized in the way that absolute value of Apollo data’s maximum yields always 1.

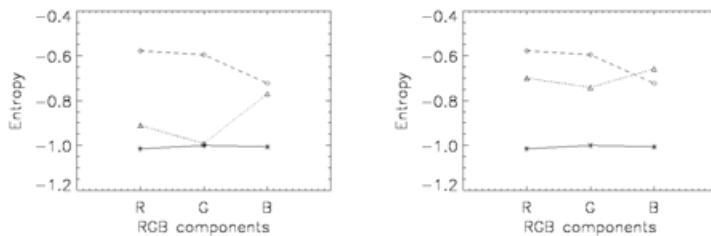


Figure 12. Normalized Burg’s entropy H_B of noise affected case. Solid line is the original Apollo image, dashed – convolved data, dotted the deconvolved one, MEM, left and ML, right. All data is normalized in the way that absolute value of Apollo data’s maximum yields always 1.

in opposition to Fourier basis. The strength of wavelets comes also from their ability to express the signal or image both in space and frequency domain at the same moment.

A very helpful feature of wavelet based image decomposition is the ability to measure an image multiscale entropy. Such measure provides not only the estimate of image information content but also a very practical estimation of this observable on different scales. As the important information in the image has always multiscale character the mutual dependences between scales are very useful in image restoration, reconstruction and analysis both in classical or Bayesian case. Hence, we also estimate the multiscale flow of information through the ideal interferometer. The formulation of multiscale entropy

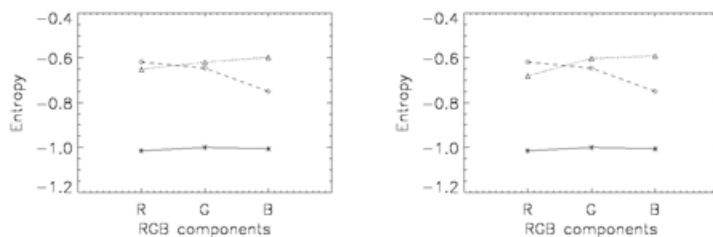


Figure 13. Normalized Burg's entropy H_B of noise affected case. Solid line is the original Apollo image, dashed – convolved data, dotted the deconvolved one, MEM, left and ML, right. All data is normalized in the way that absolute value of Apollo data's maximum yields always 1.

express the formulae:

$$H(X) = \sum_{j=1}^l \sum_{k=1}^{N_j} h(w_{j,k}) \quad (2.3)$$

where $h(w_j = \ln(p(w_j(k))))$. In order to study the behavior of information at a given scale we use mean entropy vector defined as:

$$E(j) = \frac{1}{N} \sum_{k=1}^N h(w_{j,k}). \quad (2.4)$$

It should be stressed that in all cases we are not normalizing the above vector by corresponding noise models. All simulation results were deconvolved with PSF by Maximum Likelihood (ML) and independently by Maximum Entropy (MEM) algorithms.

3. Results

3.1. Noise free case

The simulation of a perfect observing run with nonrotating interferometer yields the image shown on Fig. 4, left. There is no noise added on each level of simulation. For once established setup we have performed three observations with changing light wavelength λ . Due to low resolution coming from poor $((u, v))$ coverage the details of planet surface are hardly resolved. The information flow during noise free imaging is shown on Fig. 6. This picture illustrates multiscale entropy derived after application of ML or MEM algorithm compared to original image. The behavior of mean entropy at different scales is shown on Fig. 7.

3.2. Realistic noisy imaging

The noisy case is shown on Fig. 4, right. During simulation we have added noise according to given signal-to-noise ration (SNR) in the same way as Böker & Allen (1999). SNR was set to 10. Then the data was processed by inverse FFT what yielded the simulated observation output, Fig. 4. The image in each particular band was then deconvolved by MEM or ML. The results for both technics are shown on Fig. 5. Then we compared multiscale entropies obtained from MEM and ML as seen on Fig. 8 and their mean vectors Fig. 9.

4. Discussion

4.1. Classical imaging

The term “classical imaging” of exoplanet describes well what we really do through this work. It means the observation and “real” data reconstruction from individual “observed” images on particular wavelengths. Even if we apply the Bayesian inference by means of MEM, ML or CLEAN, usually used in interferometric data reconstruction, the entire process should be classified as classical due to no prior knowledge on an observed object. Our results suggest that in general the interferometer lowers H_S (or increase H_B). It can be noticed that MEM (Figs. 10 and 11) is recovering the data to the original data’s entropy level introducing some variation between components while ML reconstruction is oscillating around the level of interferometric data entropy. We see on multiscale entropy plots, Figs. 8 and 6 the little difference of H after processing data by an array although it rises up significantly after reconstruction. After reconstruction mean entropy vector E_j reveals significant loss of entropy on higher scales, responsible for details.

However, we have to keep in mind the significant simplifications coming from “naive” design of ideal interferometer and the assumed face-on position of planet that never takes place in real observation where the terminator must be seen.

The information processing in the procedure of interferometric imaging is crucial to understand where the information is being dissipated. Then an effective algorithm of its recovering may be constructed, knowing the precise reasons and the rate of the information leakage, like the importance of noise influence, PSF variations, disturbing signals like zodiacal light or star leakage, unknown error budget, etc. Thus not only the high-level interferometer’s model is necessary, but also the model of information processing through the instrument as well as its precise definition. Such model would help in optimization procedure of final instrument design and would lead to support data recovering algorithms. Usually, the initial data is able to be reconstructed from very ill-posed final data taking advantage of some known or predicted structures and knowledge coming from other sources or modalities. In the case of interferometric imaging the other modalities are different wavelength observations. As they all combine the real image being similar in information content on all scales they may serve as a basis for novel Bayesian strategy recovering the data due to the high correlation between different modalities.

4.2. Bayesian imaging

The possible Bayesian inference applied to exoplanet imaging can be based on the estimated level of compression of the obtained multichannel data. It is usually the multiscale measure of information content of the image data. Some prior knowledge coming from distant wavelength observation that are going to be done in advance to optical imaging, i.e. those in FIR or NIR, are also going to provide some spectroscopic information that corresponds to some surface features like green or blue spots, clouds, etc. The high level of correlation between multichannel (optical) data and that come from other modalities may serve as a basement for required methods (Molina, R. & Mateos, J. (1997)). The optimization of all accessible data and the careful analysis of its information content may lead to the Bayesian, most probable model based on the set of the most probable values of parameters. Due to rather poor present understanding of space borne and placed real interferometer, the proper design of its performance as a *noisy channel* that processes, transmits and loses the information is not known yet but should be important. The first, poorly resolved real observations of exoplanet, will also provide the basis or averaged data for optimisation of such future Bayesian algorithms.

5. Conclusions

We have discussed the basic concepts of information transfer through the interferometer. There is a leakage of information on higher scales of wavelet decomposed data. More advanced study are necessary to understand the ways of further improvements of information restoration process, both in classical and Bayesian way. The latter seems to be the proper and promising way for future interferometric imaging of exoplanets.

Acknowledgements

The author would like to acknowledge the kind and extremely helpful support from IAU to participate in this IAU Colloquium. The support made attending the conference possible.

References

- Absil, O. 2001, *Diploma thesis*, University of Liege
Angel, J. & Woolf, N. 1986, *Nature* 322, 341
Böker, T. & Allen, R. J. 1999, *ApJ* 125, 123
Bracewell, R. 1978, *Nature* 274, 780
Des Marais *et al.* 2002, *Astrobiology* 2, 153
Labeyrie, A. 1995, *Astron. Astrophys. Suppl. Ser.* 118, 517
Léger, A. 1999, *Astron. Astrophys.* 341, 304
Molina, R. *et al.* 1992, *Astron. Journ.* 104, 1662
Molina, R. & Mateos, J. 1997, *Vistas in Astronomy* 41, 373
Shanon, C. E. 1948, *Bell System Technical Journal* 27, 379
Woolf, N. & Angel, J. 1998, *Annu. Rev. Astron. Astrophysics* 36, 507