Multiscale Decomposition of Astronomical Maps: A Constrained Diffusion Method

Guang-Xing Li
South-Western Institute for Astronomy Research, Yunnan University, Kunming, 650500 Yunnan, People’s Republic of China; gxli@ynu.edu.cn
Received 2021 March 9; revised 2022 January 1; accepted 2022 January 13; published 2022 April 7

Abstract

We propose a new, efficient multiscale method to decompose a map (or signal in general) into component maps that contain structures of different sizes. In the widely used wave transform, artifacts containing negative values arise around regions with sharp transitions due to the application of band-limited filters. In our approach, the decomposition is achieved by solving a modified, nonlinear version of the diffusion equation. This is inspired by the anisotropic diffusion methods, which establish the link between image filtering and partial differential equations. In our case, the artifact issue is addressed where the positivity of the decomposed images is guaranteed.

Our new method is particularly suitable for signals which contain localized, nonlinear features, as typical of astronomical observations. It can be used to study the multiscale structures of astronomical maps quantitatively and should be useful in observation-related tasks such as background removal. We thus propose a new measure called the “scale spectrum,” which describes how the image values distribute among different components in the scale space, to describe maps. The method allows for input arrays of an arbitrary number of dimensions, and a python implementation of the algorithms is included in the Appendix and available at https://github.com/gxli/constrained_diffusion_decomposition.

Unified Astronomy Thesaurus concepts: Interstellar clouds (834); Smoothing (1929); Computational methods (1965); Astronomical methods (1043)

1. Introduction

Many astrophysical processes are multiscale in nature, and many astronomical observations generate images that contain structures at a variety of scales. Being able to decompose such images/data into components of multiple scales in a robust fashion would enable new approaches to data analysis and interpretation. For example, such a decomposition would allow the study of the relative importance of structures of different sizes. Multiscale decomposition can also be useful for observation-related tasks such as background removal.

One commonly used decomposition method is the wavelet transform (e.g., Murtagh & Starck 2002; Romeo et al. 2003). In this method, the decomposition is achieved by projecting an image onto orthonormal series generated by a wavelet, where each projection would generate an image that contains structures of a particular size. Despite its success, in the wavelet transform the decomposed maps often contain artifacts in the vicinities of regions of large contrasts, caused by the convolution with band-limited filters (Coifman & Donoho 1995). This is particularly problematic in astronomical problems, because the signals, such as the light from the sky or surface density distribution of astronomical objects, often exhibit significant, localized variations. Besides, in most astronomical or physical applications, the maps usually describe the distribution of emission from the sky or the distribution of matter or energy in space, and physically meaningful maps should contain positive values. The maps generated by the wavelet transform usually contain negative values, which are considered unphysical in many cases.

Methods like the multiscale median transform have been developed and used (Starck et al. 1998; Belloche et al. 2011). Although some of these problems are addressed, the positivity of the results is not guaranteed. Li & Burkert (2017) proposed an iterative method where the positivity is guaranteed. However, the algorithm is difficult to implement and the decomposed maps contain some artificial discontinuities which are undesirable. In this paper, inspired by a conceptual framework called the “anisotropic diffusion” (Perona & Malik 1990; Weickert 1996), we propose a new method to perform such decompositions. The method is robust and the positivity of the results is guaranteed, and can be modified to achieve decompositions that suit different purposes.

2. Constrained Diffusion Method

In the wavelet transformation, the image is convolved with a set of Gaussian kernels of increasing size, and the decomposed maps are obtained by subtracting these smoothed images from one another. Similarly, in our method, given an input map \( I(x, y) \), we produce a set of smoothed maps \( I_s(x, y) \), and the decomposed maps will be constructed by subtracting the smoothed maps from one another. In the wavelet transform, the smoothing is achieved by convolving with a Gaussian kernel, and our major improvement is to perform this by solving Equation (3), which is a modified, nonlinear version of the diffusion equation. A flowchart is presented in Figure 1. We note that in our method, the term “size” corresponds to the dispersion of the Gaussian function.

2.1. Diffusion-equation Approach to Image Processing

In image processing and computer vision, there exists a deep conceptual link between image filtering (such as smoothing) and a partial differential equation (PDE). Convolving an image with a Gaussian kernel of dispersion \( \sigma \) can be achieved by solving

\[
\frac{\partial I(x, y; t)}{\partial t} = \nabla^2 I(x, y; t),
\]

(1)
where \( I(x, y; 0) = I(x, y) \), and a proper smoothing will be achieved when \( t = l^2/2 \) where \( l \) is the smoothing length measured in terms of the dispersion.\(^1\)

Replacing Gaussian convolution with solving Equation (1) might seem to be an unnecessary complication, as Gaussian convolution is much easier to implement than solving a PDE. On a deeper level, the conceptual link between image convolution/fittering and solving PDEs has opened up the possibilities where one can develop new image operations by providing the corresponding PDEs. In the past, this had led to the proposal of a new group of techniques called anisotropic diffusion (Perona & Malik 1990; Weickert 1996), where specially designed PDEs are used to achieve feature-preserving denoising. In this paper, we import this concept where we proposed to solve a modified version of the diffusion equation to achieve multiscale image decomposition.

2.2. Constrained Diffusion

We propose to solve the following equation:

\[
\frac{\partial I_t(x, y; t)}{\partial t} = \mathcal{H}(-\nabla^2 I_t(x, y; t)) \nabla^2 I_t(x, y; t),
\]

where \( t = l^2/2 \) is “effective time”, and is related to the scale of interest \( l \) by \( t \sim l^2 \), and \( \mathcal{H} \) is the Heaviside step function (the value of \( \mathcal{H} \) is zero for negative arguments and one for positive arguments). The equation resembles a diffusion equation, with the only difference lying in the introduction of the one additional Heaviside function. It is the introduction of the Heaviside step function \( \mathcal{H}(-\nabla^2 I(x, y; t)) \) that ensures that \( I_t(x, y; t) \) decreases monotonically with \( t \), and this ensures the positivity of the decomposed maps.

Equation (2) works for input images which contain only positive values. To deal with images which also contain negative values, we further modify Equation (2) into

\[
\frac{\partial I_t}{\partial t} = \text{sgn}(I_t)\mathcal{H}(-\text{sgn}(I_t)\nabla^2 I_t)\nabla^2 I_t,
\]

where \( \text{sgn} \) is the sign function (the mathematical function that extracts the sign of a real number). We insert two \( \text{sgn}(I_t) \) functions to ensure that the results will be invariant under \( I_t \rightarrow -I_t \). This guarantees that the regions with negative values will be processed in a self-consistent way.

2.3. Decomposed Maps

After obtaining the smoothed maps, the decomposed maps are derived by subtracting them against each other, e.g., \( C_{l\rightarrow l'} = I_l(x, y) - I_{l'}(x, y) \) would contain structures with sizes larger than \( l \) yet smaller than \( l' \). As \( I_l(x, y) \) decreases monotonically with increasing scale \( l \), all of the decomposed maps \( C \) are positive by construction.
3. Implementation and Availability

3.1. Overview of the Procedure

Our algorithm can be implemented for data with an arbitrary number of dimensions. For simplicity, we consider the decomposition of a 2D map $I(x, y)$ of size $(n_x, n_y)$.

1. Define a set of scales ranges upon which the image will be decomposed. For simplicity, we use $l = 2^n$ where $n = 1, 2, 3, \ldots, n_{\text{max}}$ and $2^n_{\text{max}} \leq \min(n_x, n_y)$.

2. Take $I(x, y)$ as the input, and evolve Equation (3) from $t = 1$ to $t = l_{\text{max}}^2/2$, where $l_{\text{max}} = 2^n_{\text{max}}$. Register the values of $I_n = I(x, y; t_n)$ where $t_n = 2^{2n-1}$. In the end, we have $I_0$, which is the original image, and $n_{\text{max}}$ smoothed images.

3. Generate the decomposed maps using $C_n = I_n - I_{n-1}$. $C_n$ represents structures whose sizes are larger than $l = 2^n$ yet smaller than $l = 2^{n+1}$.

After the decomposition procedure, we obtain the output, which is a set of images of the same size $(n_x, n_y)$. Note that by summing up the decomposed maps, the original input map can be recovered, e.g., $I(x, y) = \sum_n C_n(x, y)$.

3.2. Computation of Smoothed Images

To generate the smoothed maps, we use a method we call “constrained convolution.” Our aim is to evolve Equation (3) from $t = t_1$ to $t = t_2$, where we wish to achieve this in $n$ steps. In each of our time steps, starting from $I_0$, we first compute the

Figure 2. Comparison between the wavelet transform and our constrained diffusion. Top panel: Test data. It consists of two Gaussian functions. Their dispersions are 4 and 14 pixels, respectively. Middle panel: Results from the constrained diffusion. Channel $n$ contains structures of sizes larger than $l = 2^n$ pixels yet smaller than $l = 2^{n+1}$ pixels. Bottom panel: results from the wavelet transform.
From equation (4), we have the positive and negative parts

\[ I_{h+\delta t}^{+} = \min(I_{h+\delta t}^{\text{unconstrained}}, I_{h}), \quad \text{for regions where } I_{h} > 0, \]  

\[ I_{h+\delta t}^{-} = \max(I_{h+\delta t}^{\text{unconstrained}}, I_{h}), \quad \text{for regions where } I_{h} < 0. \]  

Since our approach is effectively the finite difference theme where the truncation error is proportional to the step size (Press et al. 1992), the time step should be chosen such that \( \delta t / \delta \sim \varepsilon_{\text{rel}} \), where \( \varepsilon_{\text{rel}} \ll 1 \) is the desired accuracy. We adopt \( \varepsilon_{\text{rel}} = 0.01 \). We further require that a physically meaningful time step be smaller than yet comparable to unity, as our calculation is done on a rectangular grid and the accuracy of our calculations is limited by the discretization error. Taking these into account, the minimum time step is chosen to be 0.1. Because of this constraint, the accuracy of the first component map is limited to 0.1. A version of the code is provided in Appendix A where it has been tested using 1D, 2D, and 3D data. Due to fact that the results are invariant under \( I \rightarrow -I \), regions containing negative values will be treated consistently.

4. Performance

4.1. Comparison to the Wavelet Transform

To demonstrate the performance of our method, we first consider a simple situation where the input consists of two superimposed Gaussian functions. These Gaussian functions have dispersions of 4 and 14 pixels, respectively. Figure 2 present the results from wavelet decomposition and our constrained diffusion method. Details concerning the implementation of the wavelet transform can be found in Appendix B. From the wavelet-decomposition results, one can identify the presence of artifacts around regions of sharp transitions (e.g., regions surrounding the narrower Gaussian) caused by the application of band-limited filtering. Besides,
although the input is positive everywhere, after performing the wavelet decomposition, the decomposed maps contain coherent regions of made of negative values.

In Figure 2 we present the decomposition achieved with our constrained diffusion method. Compared to the wavelet transform, the different Gaussian components have been separated into different channels representing structures of different sizes. In contrast to the wavelet transform, the contained diffusion decomposition provides a clean separation, e.g., these Gaussian functions can be separated in such a way that they do not interfere with each other and the decomposed maps do not contain negative values. Moreover, the channel within which the structure is contained can thus be directly related to its size. E.g., the $\sigma = 4$ pixel component is most prominent in the $n = 2$ channel, which, according to our definition, contains structures of sizes $4 < l < 8$ pixels, and the $\sigma = 14$ pixel component is most prominent in $n = 3$ channel, which contains structures of sizes $8 < l < 16$ pixels.

4.2. Applications to One-, Two-, and Three-dimensional Data

Our algorithm is general and can be applied to an arbitrary number of dimensions. Some examples are presented in Figure 3.

In the 1D case, we extracted a $^{13}$CO spectrum around the central region of W43 using data from the Galactic Ring Survey data (Jackson et al. 2006), and performed a decomposition. In the results, the high-frequency noise, two narrower components, and a broad component can be separated into different channels. In the 2D case, we consider a first example where we perform decomposition to a 2D disk of a constant surface density. In the results, the edge of the disk is contained in the first channels of smaller $l$, reflecting the presence of the sharp transition at its boundary, and the middle part of the disk is separated into channels of larger $l$. We also perform a decomposition on the surface density distribution of the IC348 region using data from Zari et al. (2016). Our decomposition can separate the observed structure into different components. In the 3D case, we constructed a 3D density cube using simulation data from Clark et al. (2019). The simulations have produced a molecular cloud with a highly complex density structure, and the decomposition allows us to separate it into different components.

5. Scale Spectrum—Quantifying Multiscale Structures

The evolution of molecular clouds is a multiscale process. As a result, the density structures of molecular clouds are made of substructures that distribute over a continuous range of scales. Quantifying these multiscale density structures would provide insights into the underlying physics. Based on our decomposition, we propose a measure called the “scale spectrum” to achieve this. For a given region $R$, the scale spectrum is defined as

$$I_l = \int_R I_l(x, y) \, ds,$$

which is a representation of how matter distributes in structures of different scales. In a smooth region, matter distributes into structures of larger $l$, whereas in a structured region, matter distributes in structures of smaller $l$. Figure 4 plots the scale spectra of different subregions in the Perseus molecular cloud, where gas in, e.g., the NGC 1333 region tends to distribute in channels with smaller $l$ in comparison to other, e.g., the IC348 region. This agrees with our current understandings, as
NGC 1333 is the most evolved region in the cloud with active star formation where we expect more structures (Hatchell et al. 2005; Walawender et al. 2008). The scale spectrum offers a quantitative way to describe a region, which enables quantitative comparisons.

6. Concluding Remarks

We propose a new method to decompose, in a robust way, an image/signal into a set of maps that contain structures at different scales. The method consists of producing smoothed images using a modified, nonlinear version of the diffusion equation, and the decomposed images are obtained by subtracting these smoothed images from one another. Compared to the widely used wavelet transform, our method is robust against the presence of sharp transitions. If the input image is positive, the positivity of the decomposed maps is guaranteed. The positivity is often necessary or desirable for subsequent analyses and interpretations. The method is implemented as a python3 program available in Appendix A. A maintained version is available through https://github.com/gxli/constrained_diffusion_decomposition. The method is applicable for \( n \)-dimensional data. We expect our decomposition to be useful in analyzing and quantifying images, and in practical tasks such as data visualization and background removal.

This algorithm allows us to propose a new measure called the “scale spectrum” to quantify images. It describes how the observed values distribute in substructures of different sizes, leading to a new, quantitative way to describe and compare structures. This paper also introduces the anisotropic diffusion method to the astronomical community. In this new approach, decomposition is achieved by solving a modified diffusion equation. This conceptual link should enable other new, meaningful ways to decompose maps by providing similarly modified PDEs to achieve different purposes.

G.-X.L. acknowledges support from NSFC grants W820301904 and 12033005.
Appendix A
Implementation and Usage

A python3 implementation of our method can be found below (constrained_diffusion_decomposition.py).

```python
#!/usr/bin/env python
from astropy.io import fits
import sys
import numpy as np
from math import log
from scipy import ndimage

def constrained_diffusion_decomposition(data,
    e_rel=3e-2,
    max_n=None, sm_mode='reflect',):

    
    """
    perform constrained diffusion decomposition
    inputs:
    data: n-dimensional array
    e_rel: relative error, a smaller e_rel means a better
           accuracy yet a larger computational cost
    max_n: maximum number of channels. Channel number
           ranges from 0 to max_n
           if None, the program will calculate it automatically
    sm_mode: {'reflect', 'constant', 'nearest', 'mirror', 'wrap'}, optional
             The mode parameter determines how the input array is extended beyond its
             boundaries in the convolution operation. Default is 'reflect'.
    output:
    results: of constrained diffusion decomposition. Assuming that the input
             is a n-dimensional array, then the output would be a n+1 dimensional
             array. The added dimension is the scale. Component maps can be accessed
             via output[n], where n is the channel number.
             output[1] contains structures of sizes larger than 2**i pixels
             yet smaller than 2**(i+1) pixels.
             residual: structures too large to be contained in the results
    """

    ntot = int(log(min(data.shape))/log(2) - 1)
    # the total number of scale map

    result = []
    # residual = []
    # residual.append(data)
    if max_n is not None:
        ntot = np.min(ntot, max_n)
        print("ntot", ntot)

    diff_image = data.copy() * 0

    for i in range(ntot):
        print("i = ", i)
        channel_image = data.copy() * 0

        # computing the step size
        scale_end = float(pow(2, i + 1))
        scale_beginning = float(pow(2, i))
        t_end = scale_end**2 / 2  # t at the end of this scale
        t_beginning = scale_beginning**2 / 2 # t at the beginning of this scale

        if i == 0:
            delta_t_max = t_beginning * 0.1
        else:
            delta_t_max = t_beginning * e_rel
```

7
and a maintained version is available at https://gxli.github.io/Constrained-Diffusion-Decomposition/. A copy of these files have also be deposited to Zenodo at doi:10.5281/zenodo.6375571. The code can be used as a python package

```python
niter = int((t.end - t.beginning) / delta_t_max + 0.5)
delta_t = (t.end - t.beginning) / niter
kernel_size = np.sqrt(2 * delta_t)  # size of gaussian kernel

print("kernel_size", kernel_size)
for kk in range(niter):
    smooth_image = ndimage.gaussian_filter(data, kernel_size,
                                           mode=sm_mode)
    sm_image_1 = np.minimum(data, smooth_image)
    sm_image_2 = np.maximum(data, smooth_image)

    diff_image_1 = data - sm_image_1
    diff_image_2 = data - sm_image_2

    diff_image = diff_image * 0
    positions_1 = np.where(np.logical_and(diff_image_1 > 0, data > 0))
    positions_2 = np.where(np.logical_and(diff_image_2 < 0, data < 0))

    diff_image[positions_1] = diff_image_1[positions_1]
    diff_image[positions_2] = diff_image_2[positions_2]

    channel_image = channel_image + diff_image

    data = data - diff_image
    # data = ndimage.gaussian_filter(data, kernel_size)  # !!!!
    result.append(channel_image)
    residual.append(data)

residual = data
return result, residual
```

and as a shell script

```bash
# assuming that data is a N-dimensional array
import constrained_diffusion_decomposition as cdd
# the decomposition can be computed via
result, residuals = cdd.constrained_diffusion_decomposition(data)
```

Given a input $I$ of size $(n_x, n_y)$, the output $C$ takes the form of $(n_i, n_x, n_y)$ where decomposed maps are arranged such that $C[i]$ has same size as the input, and it contain structures of sizes than range from $2^i$ to $2^{i+1}$ pixels. The “size” corresponds to the dispersion of the a Gaussian function. The residual contains the residuals, e.g., structures whose sizes are even larger than $2^{i+1}$ pixels. The program has been tested with one-, two-, and three-dimensional data.
Appendix B
The Wavelet Transform

We compare our method against a version of the wavelet transform called the à Trous wavelet transform (Starck & Murtagh 2008). The wavelet decomposition is a procedure where the decomposition is achieved by projecting an image onto a set of Hilbert bases, where different bases correspond to structures at different scales. In a simple case, the wavelet transform can be realized by a set of convolutions where the wavelet component $n$ is defined as

$$I_n(x, y) = I(x, y) * (g(2^n) - g(2^{n-1})),$$

where $I(x, y)$ represents the original emission map and $g(x)$ represents the Gaussian function where $x$ is the dispersion.

Appendix C
Performance Demonstration

In Figures 5 and 6 we apply our constrained diffusion decomposition to a few examples, and compare the results with those from the wavelet decompositions.

**Figure 5.** Comparison between the results from our constrained diffusion decomposition (upper panels) to those from the wavelet transformation (lower panels). These maps are derived by applying these methods to a surface density map of the IC348 region obtained from Zari et al. (2016).
Figure 6. Comparison between the results from our constrained diffusion decomposition (upper panels) to those from the wavelet transformation (lower panels). The results are derived by applying these methods to a map containing a disk of a constant surface density.
References

Belloche, A., Schuller, F., Parise, B., et al. 2011, A&A, 527, A145
Clark, P. C., Glover, S. C. O., Ragan, S. E., & Duarte-Cabral, A. 2019, MNRAS, 486, 4622
Coifman, R. R., & Donoho, D. 1995, Translation-Invariant De-Noising (Berlin: Springer-Verlag), 125
Hatchell, J., Richer, J. S., Fuller, G. A., et al. 2005, A&A, 440, 151
Jackson, J. M., Rathborne, J. M., Shah, R. Y., et al. 2006, ApJS, 163, 145
Li, G.-X., & Burkert, A. 2017, MNRAS, 464, 4096
Murtagh, F., & Starck, J.-L. 2002, in Wavelets and Multiscale Transform in Astronomical Image Processing, ed. J. Abello, P. M. Pardalos, & M. G. C. Resende (Boston, MA: Springer US), 473
Perona, P., & Malik, J. 1990, ITPAM, 12, 629
Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. 1992, Numerical Recipes in FORTRAN. The Art of Scientific Computing (Cambridge: Cambridge Univ. Press)
Romeo, A. B., Horellou, C., & Bergh, J. 2003, MNRAS, 342, 337
Starck, J., & Murtagh, F. 2008, Astronomical Image and Data Analysis (Berlin: Springer)
Starck, J.-L., Murtagh, F. D., & Bijaoui, A. 1998, Image and Data Analysis: The Multiscale Approach (Cambridge: Cambridge Univ. Press)
Walawender, J., Bally, J., Francesco, J. D., Jørgensen, J., & Getman, K. 2008, in Handbook of Star Forming Regions, Volume I: The Northern Sky, ed. B. Reipurth (San Francisco, CA: ASP), 346
Weickert, J. 1996, Anisotropic Diffusion In Image Processing (Berlin: Vieweg +Teubner Verlag)
Zari, E., Lombardi, M., Alves, J., Lada, C. J., & Bouy, H. 2016, A&A, 587, A106