# Link from "movie night" worldwidetelescope.org



# Astronomy as

### Alyssa A. Goodman

Professor of Astronomy & Founding Director of the Initiative in Innovative Computing, Harvard University

Scholar-in-Residence, WGBH Boston



## Science + Technology

#### LETTERS

#### NATURE Vol 457 1 January 2009



Figure 2 Comparison of the 'dendrogram' and 'CLUMPFIND' featureidentification algorithms as applied to <sup>13</sup>CO emission from the L1448 region of Perseus. a, 3D visualization of the surfaces indicated by colours in the dendrogram shown in c. Purple illustrates the smallest scale selfgravitating structures in the region corresponding to the leaves of the dendrogram; pink shows the smallest surfaces that contain distinct selfgravitating leaves within them; and green corresponds to the surface in the data cube containing all the significant emission. Dendrogram branches corresponding to self-gravitating objects have been highlighted in yellow over the range of  $T_{\rm mb}$  (main-beam temperature) test-level values for which the virial parameter is less than 2. The x-y locations of the four 'selfgravitating' leaves labelled with billiard balls are the same as those shown in Fig. 1. The 3D visualizations show position-position-velocity (p-p-v) space. RA, right ascension; dec., declination. For comparison with the ability of dendrograms (c) to track hierarchical structure, d shows a pseudodendrogram of the CLUMPFIND segmentation (b), with the same four labels used in Fig. 1 and in a. As 'clumps' are not allowed to belong to larger structures, each pseudo-branch in **d** is simply a series of lines connecting the maximum emission value in each clump to the threshold value. A very large number of clumps appears in b because of the sensitivity of CLUMPFIND to noise and small-scale structure in the data. In the online PDF version, the 3D cubes (**a** and **b**) can be rotated to any orientation, and surfaces can be turned on and off (interaction requires Adobe Acrobat version 7.0.8 or higher). In the printed version, the front face of each 3D cube (the 'home' view in the interactive online version) corresponds exactly to the patch of sky shown in Fig. 1, and velocity with respect to the Local Standard of Rest increases from front  $(-0.5 \text{ km s}^{-1})$  to back  $(8 \text{ km s}^{-1})$ .

data, CLUMPFIND typically finds features on a limited range of scales, above but close to the physical resolution of the data, and its results can be overly dependent on input parameters. By tuning CLUMPFIND's two free parameters, the same molecular-line data set<sup>8</sup> can be used to show either that the frequency distribution of clump mass is the same as the initial mass function of stars or that it follows the much shallower mass function associated with large-scale molecular clouds (Supplementary Fig. 1).

Four years before the advent of CLUMPFIND, 'structure trees'<sup>9</sup> were proposed as a way to characterize clouds' hierarchical structure

64

©2009 Macmillan Publishers Limited. All rights reserved

using 2D maps of column density. With this early 2D work as inspiration, we have developed a structure-identification algorithm that abstracts the hierarchical structure of a 3D (p-p-v) data cube into an easily visualized representation called a 'dendrogram'<sup>10</sup>. Although well developed in other data-intensive fields<sup>11,12</sup>, it is curious that the application of tree methodologies so far in astrophysics has been rare, and almost exclusively within the area of galaxy evolution, where 'merger trees' are being used with increasing frequency<sup>13</sup>.

Figure 3 and its legend explain the construction of dendrograms schematically. The dendrogram quantifies how and where local maxima of emission merge with each other, and its implementation is explained in Supplementary Methods. Critically, the dendrogram is determined almost entirely by the data itself, and it has negligible sensitivity to algorithm parameters. To make graphical presentation possible on paper and 2D screens, we 'flatten' the dendrograms of 3D data (see Fig. 3 and its legend), by sorting their 'branches' to not cross, which eliminates dimensional information on the *x* axis while preserving all information about connectivity and hierarchy. Numbered 'billiard ball' labels in the figures let the reader match features between a 2D map (Fig. 1), an interactive 3D map (Fig. 2a online) and a sorted dendrogram (Fig. 2c).

A dendrogram of a spectral-line data cube allows for the estimation of key physical properties associated with volumes bounded by isosurfaces, such as radius (R), velocity dispersion ( $\sigma_v$ ) and luminosity (L). The volumes can have any shape, and in other work<sup>14</sup> we focus on the significance of the especially elongated features seen in L1448 (Fig. 2a). The luminosity is an approximate proxy for mass, such that  $M_{\text{lum}} = X_{13\text{CO}}L_{13\text{CO}}$ , where  $X_{13\text{CO}} = 8.0 \times 10^{20} \text{ cm}^2 \text{ K}^{-1} \text{ km}^{-1} \text{ s}$ (ref. 15; see Supplementary Methods and Supplementary Fig. 2). The derived values for size, mass and velocity dispersion can then be used to estimate the role of self-gravity at each point in the hierarchy, via calculation of an 'observed' virial parameter,  $\alpha_{obs} = 5\sigma_v^2 R/GM_{lum}$ . In principle, extended portions of the tree (Fig. 2, yellow highlighting) where  $\alpha_{obs} < 2$  (where gravitational energy is comparable to or larger than kinetic energy) correspond to regions of p-p-v space where selfgravity is significant. As  $\alpha_{obs}$  only represents the ratio of kinetic energy to gravitational energy at one point in time, and does not explicitly capture external over-pressure and/or magnetic fields16, its measured value should only be used as a guide to the longevity (boundedness) of any particular feature.



Figure 3 | Schematic illustration of the dendrogram process. Shown is the construction of a dendrogram from a hypothetical one-dimensional emission profile (black). The dendrogram (blue) can be constructed by 'dropping' a test constant emission level (purple) from above in tiny steps (exaggerated in size here, light lines) until all the local maxima and mergers are found, and connected as shown. The intersection of a test level with the emission is a set of points (for example the light purple dots) in one dimensions. The dendrogram of 3D data shown in Fig. 2c is the direct analogue of the tree shown here, only constructed from 'isosurface' rather than 'point' intersections. It has been sorted and flattened for representation on a flat page, as fully representing dendrograms for 3D data cubes would require four dimensions.



### The most incomprehensible

# "Astronomy as I See It"



- Albert Einstein

# **Relative Strengths**



# "Interocularity"

## (see work of John Tukey

# "Image and Meaning"

(see work of Felice Frankel, and imageandmeaning.org)

# Are we held back by confining tecnological tools?





Autog. par Regnier, 8. Pas. Ste Marie St Gate à Paris

# Galileo, Science & Technology

ler Prinke Yolike Galily Humilin " Serve Della Ser" V. inuigilan To assiduance at to ogny phinic & borere to solar satisfas alianis the nene della caum to Max umation nello fre Die Di Padona. Invers Dawers determinate & projentare al Joy Je [ Dichiele it I give Di Finamente inghins bile & ogs regois et in irea maritima o terrestre sino Ditenere que to nave attifizione (maggior jeg to at where a Differentione A o for I Valiale anoto Talle più & Sik speculazioni & pro bettina na quantaggio di poprise Legnice Vele dell' in mia Frae here it put i singe frima & get junopra noi et Distingues I me men et la qualité Sei Vattelly quichiare le sue forme ballestivitiatta carcia al amostomento o alla fuga, o pure and nella capaqua apirta sidere et partivlarry Distingutre agri su ind to et prepitamento crating treto it no retor in talk within Line . i meglis whe Adi 14 è may X \* la profi stante salla 3ª 1 Poppie Ta La spatio Delle 3 our erati and com maggine Del Dinastro Do 7 et e mass in lines rates .

1	and the second	Rent Contraction
7	* •0 *	h * 0
8	0***	* 0
ø	* * O	9.0.
H	** 0	4 * •() * *
12	× •() *	<b>u</b> .0. '0. '0.
13.	+ 0***	<b>ы</b> О
ış	0 • • • •	(72 · 0····
1¢	o	Jr O.
16	****	14 0 . 13 0
17	*•	24 0 24

JIDE LEUS NUNCIUS

On the third, at the seventh hour, the stars were arranged in this quence. The eastern one was 1 minute, 30 seconds from Jupiter 2 closest western one 2 minutes; and the other western one wa

o minutes removed from this one. They were absolutely on the ame straight line and of equal magnitude.

On the fourth, at the second hour, there were four stars arour upiter, two to the east and two to the west, and arranged precise

\* \* **O** \* \* We

on a straight line, as in the adjoining figure. The easternmost wa listant 3 minutes from the next one, while this one was 40 second rom Jupiter; Jupiter was 4 minutes from the nearest western one. d this one 6 minutes from the westernmost one. Their magnitude, ere nearly equal; the one closest to Jupiter appeared a little smaller ian the rest. But at the seventh hour the eastern stars were only o seconds apart. Jupiter was 2 minutes from the nearer eastern

\*\* O \* \* West

one, while he was 4 minutes from the next western one, and this one was 3 minutes from the westernmost one. They were all equal and extended on the same straight line along the ecliptic.

On the fifth, the sky was cloudy.

East

East

On the sixth, only two stars appeared flanking Jupiter, as is seen

West

in the adjoining figure. The eastern one was 2 minutes and the vestern one 3 minutes from Jupiter. They were on the same straight line with Jupiter and equal in magnitude.

On the seventh, two stars stood near Jupiter, both to the east

Notes for & re-productions of Siderius Nuncius (1610)

# Star (and Planet, and Moon) Formation 101



## "IMF"? "CMF"?

Note: IMF= "Initial Mass Funciton" of Stars, not "International Monetary Fund."



Alves, Lombardi & Lada 2007



BUT: Beautiful images like this do not reveal *internal* structure directly...

simulations



>2D observations

# Astronomical Medicine

Alyssa Goodman (IIC/CfA/FAS) Michael Halle (IIC/SPL/HMS) Ron Kikinis (SPL/HMS) Douglas Alan (IIC) Michelle Borkin (FAS/IIC) Jens Kauffmann (CfA/IIC) Erik Rosolowsky (CfA/UBC Okanagan) Nick Holliman (U. Durham)



# The Astronomical Medicine Story



### The size Solve A PLETE = COordinated Molecular VIL: 63 PW 666 Line Exinction Thermal Emission

### mm peak (Enoch et al. 2006)

sub-mm peak (Hatchell et al. 2005, Kirk et al. 2006)

<sup>13</sup>CO (Ridge et al. 2006)

mid-IR IRAC composite from c2d data (Foster, Laakso, Ridge, et al. in prep.)

Optical image (Barnard 1927)

om: 227% Angle: 0

n: 17249

# "Three" Dimensions: Spectral-Line Mapping

We wish we could measure...

But we can measure...





## Radio Spectral-line Observations of Interstellar Clouds



Wednesday, June 17, 2009

## Radio Spectral-line Observations of Interstellar Clouds



Alves, Lada & Lada 1999

# Velocity as a "Fourth" Dimension



## Astronomical Visualization Tools are Traditionally 2D



### "3D"=movies

## "Astronomical Medicine"



"z" is depth into head

"z" is line-of-sight velocity



(This kind of "series of 2D slices view" is known in the Viz as "the grand tour")



...and the science is in the interpretation of these measurements into physical quantities <u>& processes</u>.



Image of NGC1333

# What can we observe?



X-Ray of Human Skull, c. 1920

# What can we observe?



# What can we observe? $(\vec{x}, t)$ I(Time Intensity Spatial Position (x, y, z)**Cardiac Motion**

# What can we observe?



Optical (B,V,R) image of NGC1333



Human Ear, Thermal Infrared 100.7°F

100

90

80

76.2°F

# What can we observe?





### COMPLETE Perseus

/iew size: 1305 × 733 /L: 63 WW: 127

### mm peak (Enoch et al. 2006)

sub-mm peak (Hatchell et al. 2005, Kirk et al. 2006)

<sup>13</sup>CO (Ridge et al. 2006)

mid-IR IRAC composite from c2d data (Foster, Laakso, Ridge, et al. in prep.)

Optical image (Barnard 1927)

i: 1/249 om: 227% Angle: 0



3D Viz made with VolView

# AstronomicalMedicine@



# Real 3D space



Head "x"



3D rendering: <u>GE Healthcare</u>

## "Position-Position-Velocity" Space



Sky "x" (Right Ascension)



3D rendering: AstroMed /N. Holliman (U. Durham), using VolView (ITK-based)



## "Tasting" Magnetohydrodynamic Simulations



Simulations of Bate 2009

# Tasting $I(E, s, \vec{x}, t)$



۵

# The Taste-Testing Process



# "Seeing" and "Tasting" The Role Self-Gravity in Star Formation

#### LETTERS

#### NATURE Vol 457 | 1 January 2009



identification algorithms as applied to <sup>13</sup>CO emission from the L1448 region of Perseus. a, 3D visualization of the surfaces indicated by colours in the dendrogram shown in c. Purple illustrates the smallest scale selfgravitating structures in the region corresponding to the leaves of the dendrogram; pink shows the smallest surfaces that contain distinct selfgravitating leaves within them; and green corresponds to the surface in the data cube containing all the significant emission. Dendrogram branches corresponding to self-gravitating objects have been highlighted in yellow over the range of Tmb (main-beam temperature) test-level values for which the virial parameter is less than 2. The x-y locations of the four 'selfgravitating' leaves labelled with billiard balls are the same as those shown in Fig. 1. The 3D visualizations show position-position-velocity (p-p-v) space. RA, right ascension; dec., declination. For comparison with the ability of dendrograms (c) to track hierarchical structure, d shows a pseudodendrogram of the CLUMPFIND segmentation (b), with the same four labels used in Fig. 1 and in a. As 'clumps' are not allowed to belong to larger structures, each pseudo-branch in d is simply a series of lines connecting the maximum emission value in each clump to the threshold value. A very large number of clumps appears in b because of the sensitivity of CLUMPFIND to noise and small-scale structure in the data. In the online PDF version, the 3D cubes (a and b) can be rotated to any orientation, and surfaces can be turned on and off (interaction requires Adobe Acrobat version 7.0.8 or higher). In the printed version, the front face of each 3D cube (the 'home' view in the interactive online version) corresponds exactly to the patch of sky shown in Fig. 1, and velocity with respect to the Local Standard of Rest increases from front  $(-0.5 \text{ km s}^{-1})$  to back  $(8 \text{ km s}^{-1})$ .

data, CLUMPFIND typically finds features on a limited range of scales, above but close to the physical resolution of the data, and its results can be overly dependent on input parameters. By tuning CLUMPFIND's two free parameters, the same molecular-line data set<sup>8</sup> can be used to show either that the frequency distribution of clump mass is the same as the initial mass function of stars or that it follows the much shallower mass function associated with large-scale molecular clouds (Supplementary Fig. 1).

Four years before the advent of CLUMPFIND, 'structure trees'9 were proposed as a way to characterize clouds' hierarchical structure

©2009 Macmillan Publishers Limited. All rights reserved

using 2D maps of column density. With this early 2D work as inspiration, we have developed a structure-identification algorithm that abstracts the hierarchical structure of a 3D  $(p-p-\nu)$  data cube into an easily visualized representation called a 'dendrogram'<sup>10</sup>. Although well developed in other data-intensive fields<sup>11,12</sup>, it is curious that the application of tree methodologies so far in astrophysics has been rare, and almost exclusively within the area of galaxy evolution, where 'merger trees' are being used with increasing frequency<sup>13</sup>.

Figure 3 and its legend explain the construction of dendrograms schematically. The dendrogram quantifies how and where local maxima of emission merge with each other, and its implementation is explained in Supplementary Methods. Critically, the dendrogram is determined almost entirely by the data itself, and it has negligible sensitivity to algorithm parameters. To make graphical presentation possible on paper and 2D screens, we 'flatten' the dendrograms of 3D data (see Fig. 3 and its legend), by sorting their 'branches' to not cross, which eliminates dimensional information on the x axis while preserving all information about connectivity and hierarchy. Numbered 'billiard ball' labels in the figures let the reader match features between a 2D map (Fig. 1), an interactive 3D map (Fig. 2a online) and a sorted dendrogram (Fig. 2c).

A dendrogram of a spectral-line data cube allows for the estimation of key physical properties associated with volumes bounded by isosurfaces, such as radius (R), velocity dispersion ( $\sigma_v$ ) and luminosity (L). The volumes can have any shape, and in other work14 we focus on the significance of the especially elongated features seen in L1448 (Fig. 2a). The luminosity is an approximate proxy for mass, such that  $M_{\text{lum}} = X_{13\text{CO}}L_{13\text{CO}}$ , where  $X_{13\text{CO}} = 8.0 \times 10^{20} \text{ cm}^2 \text{ K}^{-1} \text{ km}^{-1} \text{ s}$ (ref. 15; see Supplementary Methods and Supplementary Fig. 2). The derived values for size, mass and velocity dispersion can then be used to estimate the role of self-gravity at each point in the hierarchy, via calculation of an 'observed' virial parameter,  $\alpha_{obs} = 5\sigma_v^2 R/GM_{lum}$ In principle, extended portions of the tree (Fig. 2, yellow highlighting) where  $\alpha_{obs} < 2$  (where gravitational energy is comparable to or larger than kinetic energy) correspond to regions of p-p-v space where selfgravity is significant. As  $\alpha_{obs}$  only represents the ratio of kinetic energy to gravitational energy at one point in time, and does not explicitly capture external over-pressure and/or magnetic fields16, its measured value should only be used as a guide to the longevity (boundedness) of any particular feature.



Figure 3 | Schematic illustration of the dendrogram process. Shown is the construction of a dendrogram from a hypothetical one-dimensional emission profile (black). The dendrogram (blue) can be constructed by 'dropping' a test constant emission level (purple) from above in tiny steps (exaggerated in size here, light lines) until all the local maxima and mergers are found, and connected as shown. The intersection of a test level with the emission is a set of points (for example the light purple dots) in one dimension, a planar curve in two dimensions, and an isosurface in three dimensions. The dendrogram of 3D data shown in Fig. 2c is the direct than 'point' intersections. It has been sorted and flattened for representation on a flat page, as fully representing dendrograms for 3D data cubes would require four dimensions.



### Goodman et al. Nature, 2009

## Bonus: Using Extra Senses & Re-Using Software



# Astronomy as I Touch it?

home about the iic research education people events employment	Initiative in Innovative		
reaching the iic INITIATIVE IN	home > research scientists' discovery room lab (sdr lab)		
INNOVATIVE COMPUTING (IIC) NEWSLETTER Stay informed on our latest news!	Lead investigators Chia Shen (IIC ) and Hanspeter Pfister (SEAS/IIC) Project staff		
E-mail: *	Michael Horn, Meekal Bajaj, Matthew Tobiasz and Matthias Lee Description		
<ul> <li>Subscribe</li> <li>Unsubscribe</li> </ul>	The Scientists' Discovery Room (SDR) is a next-generation visual digital laboratory for science discovery, collaborative learning and education. Our research focuses on		

visual digital laboratory for science discovery, collaborativi learning and education. Our research focuses on experimenting with new modalities of human-computer interaction and visualization, to create a new genre of navigation, exploration and detailed analyses in multidimensional information spaces. All projects in SDR are in close collaboration with domain scientists and educators.

INVOLV is a generalizable multi-user interactive visualization framework for large hierarchical data sets. In this project, we address the visual layout of both the primary data representation and the overlay of alternate structures of the same data. Our first case study is the visualization of life on earth based on the Encyclopedia of Life (www.eol.org)

and Tree of Life (<u>www.tolweb.org</u>). The user interface provides free-form exploration of more than 1.2 million named species while communicating issues of biodiversity and phylogeny. The current visualization combines a Voronoi Treemap tessellation with innovative human-computer interaction designs to support collaborative exploration and learning. Please visit <u>www.involvweb.org</u> for more information on this project.

ve Computing at Harvard

CThru, a collaborative endeavor with Molecular and Cellular Biology faculty, aims to develop a selfguided learning environment. In CThru, we examine methods for constructing interactive video-based educational modules. Using the animation "The Inner Life of the Cell" as a testbed, CThru addresses research issues of embedding interactive visible objects, extensive multimedia information and manipulatable 3D models within a video flow, replacing sequential video viewing with the experience of exploring and manipulating in a multi-dimentional information space.

WeSpace is a collaborative work space that integrates a large data wall with a multi-user multi-touch table. WeSpace has been developed for a population of scientists who frequently meet in small groups for data exploration and visualization. It provides a low overhead walk-up and share environment for users with their own personal applications and laptops.

**LivOlay** is an interactive image overlay tool that enables the rapid visual overlay of live data rendered in different applications. Our tool addresses datasets in which visual registration of the information is necessary in order to allow for thorough understanding and visual analysis.

### Slideshow: Tabletop Computers Continued

By Meredith Ringel Morris

#### First Published December 2008

⊠ <u>Email</u>	A Print	Comments (1)	P Reprints	Newsletters	
Del.icio.us Digg / Slashdot					



PHOTO: HAO JIANG, DANIEL WIGDOR, CLIFTON FORLINES, AND CHIA SHEN

**UBITABLE**: Users can interact with surface computers through auxiliary devices, such as laptops, phones, and PDAs. The display on the auxiliary device can convey private or sensitive content to a single user, while group-appropriate content can appear on the tabletop display. Chia Shen and her colleagues at Mitsubishi Electric Research Laboratories, in Cambridge, Mass., have explored auxiliary interactions with surface computers in their UbiTable project, in which two people with laptops collaborate over a tabletop display. Recently, Shen expanded the UbiTable into an interactive room called the WeSpace. People can share data on their laptops with other people in the room, using both a table and a large display wall. Here, three Harvard University astrophysicists discuss radio and IR spectrum images using the WeSpace.

### http://iic.harvard.edu/research/scientists-discovery-room-lab-sdr-lab

### http://spectrum.ieee.org/dec08/6999/9

Submit

revious issues

### The Scientists' Discovery Room (Shen & Pfister)





movie courtesy Daniel Wigdor, equipment now in Chia Shen's SDR lab at SEAS

What about data visualization... ...is easier now than before? fast computation, animation, 3D

...was easier before than now? craftsmanship

...should be easier in the future?

modular craftsmanship







# I think we can aim for general, interoperable, (viz) tools with a "modular craftsmanship" approach.

### ....was easier before than now? CIVA craftsmanship

# ...should be easier in the future?

modular craftsmanship

Astronomy as I See, Touch, Taste, and Hear It

