Astronomers!

Do you know where your galaxies are?

Introducing the *Duchamp* Source Finder

Matthew Whiting, Australia Telescope National Facility, CSIRO

The Problem: Making the most of your surveys

Large-scale spectral-line radio surveys are an effective way to study populations of galaxies, masers, or sites of star-formation. Surveying on a large scale allows one to both probe to the fainter flux levels and remain sensitive to the rare but interesting objects at the brightest fluxes.

Effective use of large-scale surveys is dictated by the efficiency of detection of the objects of interest. Finding the extreme bright objects is often the easy bit – to find the more typical sources you need to push down to fainter fluxes close to the noise level. The complexity and volume of three-dimensional survey data also requires a large degree of automation and reliability to find the desired sources quickly and in a uniform way.

The Solution: The *Duchamp* Source Finder

Duchamp is our solution to the problem of three-dimensional source finding. It is a stand-alone software package, designed to work with spectral-line FITS cubes and produce source lists *and* graphical results showing detected objects.

Duchamp is optimised for the case of a large number of separated sources embedded in a cube dominated by noise – the typical situation expected for HI or maser surveys.

Innovations in Searching

A key aspect of source detection is the application of a threshold. This is calculated by using either a simple *n-sigma* cutoff, or derived through the use of the False Discovery Rate method[1],which controls the number of false detections. The statistics for the threshold calculations are obtained through robust methods, and can be measured from the full cube or a specified subset.

Searching in 3D is in general a complex problem. *Duchamp* approaches this by searching in each 2D channel map separately[2], and comparing detections made in one channel with those in neighbouring channels. An efficient merging algorithm then merges the 2D detections to form 3D objects.

Beating the noise

The limiting factor in detecting faint objects is the background noise, and *Duchamp* provides ways to minimise its effects. Simple smoothing before searching is possible, either spectrally using a Hanning filter, or spatially using a 2D Gaussian kernel.

Alternatively, the cube can be reconstructed using the *à trous* wavelet technique[3]. This filters the cube at a range of scales, thresholding each scale and only keeping those pixels with significant signal. This very effectively removes noise from the cube, allowing searching to be done on much cleaner data.

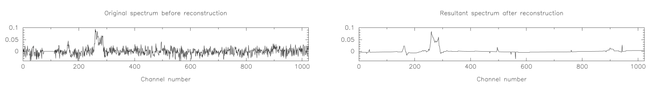

Figure 1: An example of the *à trous* algorithm used in one dimension to reconstruct a spectrum. Note the double-horned shape of the resultant main object, and the preservation of the weak features around channels 170 & 900.

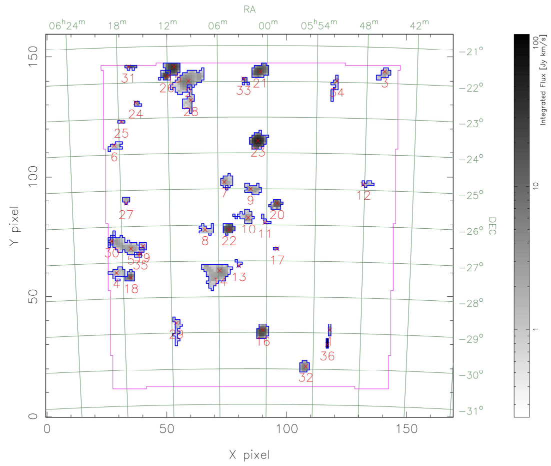

Figure 2: Example graphical output of *Duchamp* showing the zeroth moment of each detection, in its appropriate spatial location. This is provided as a postscript file and displayed in a PGPLOT window.

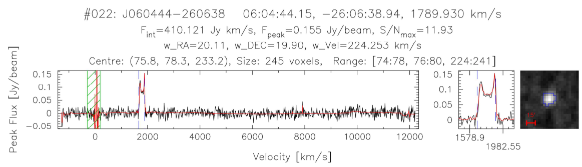

Figure 3: Example graphical output showing the spectral nature, the zeroth moment and the basic properties of a detection. A postscript file is produced showing such a plot for each object.

Obj#	Name	X	Y	Z	RA	DEC	VEL [km/s]	w_RA [arcmin]	w_DEC [arcmin]	w_VEL [km/s]	F_int [Jy km/s]	F_peak [Jy/beam]	S/Nmax	X1	X2	Y1	Y2	Z1	Z2	Npix [pix]	Flag
1	J055156-285539	118.0	35.8	73.0	05:51:56.54	-28:55:39.34	-323.437	3.58	20.10	52.765	5.131	0.123	9.53	118	118	34	38	70	74	11	
2	J060929-215751	59.0	140.3	114.6	06:09:29.23	-21:57:51.61	225.113	56.58	43.29	65.957	1467.395	0.213	16.41	52	65	135	145	113	118	169	
3	J054556-214408	140.9	143.0	114.7	05:45:56.03	-21:44:08.75	227.293	23.52	20.78	26.383	46.146	0.090	6.93	138	143	141	145	114	116	33	
4	J061839-271654	29.5	59.9	115.0	06:18:39.30	-27:16:54.35	231.217	28.86	23.09	26.383	47.693	0.107	8.25	26	32	57	62	114	116	23	E
5	J061738-263227	32.6	71.1	115.4	06:17:38.61	-26:32:27.67	236.625	61.04	30.20	79.148	712.933	0.117	9.06	26	40	68	75	113	119	110	E
6	J061827-234557	28.6	112.6	115.6	06:18:27.08	-23:45:57.80	238.274	24.51	15.25	26.383	18.895	0.070	5.38	26	31	111	114	115	117	21	
7	J060458-244700	74.9	98.2	116.3	06:04:58.37	-24:47:00.29	247.617	20.15	27.89	39.574	78.679	0.109	8.40	73	77	95	101	115	118	38	
8	J060724-260826	66.8	77.8	117.0	06:07:24.83	-26:08:26.63	256.588	24.21	19.75	26.383	35.721	0.094	7.26	64	69	76	80	116	118	24	
9	J060142-250028	86.0	94.9	118.0	06:01:42.80	-25:00:28.12	269.778	27.99	20.02	26.383	104.411	0.124	9.57	83	89	93	97	117	119	49	
10	J060230-254647	83.3	83.3	118.2	06:02:30.15	-25:46:47.13	273.178	36.04	23.94	65.957	92.467	0.118	9.10	78	86	81	86	117	122	40	
11	J060011-255152	91.1	82.0	118.4	06:00:11.53	-25:51:52.67	275.782	15.94	16.06	26.383	2.849	0.070	5.38	90	93	81	84	118	120	6	
12	J054741-244835	133.7	97.2	118.7	05:47:41.60	-24:48:35.62	279.972	19.71	12.55	26.383	9.339	0.074	5.71	132	136	96	98	118	120	14	
13	J060329-270410	80.0	64.0	120.0	06:03:29.59	-27:04:10.01	297.333	12.04	11.97	26.383	1.123	0.064	4.91	79	81	63	65	119	121	4	
14	J060613-272009	70.9	59.9	121.3	06:06:13.44	-27:20:09.07	314.462	52.44	43.59	105.531	1101.345	0.150	11.56	65	77	53	63	120	128	146	
15	J061118-213642	52.5	145.5	162.5	06:11:18.98	-21:36:42.33	857.976	28.33	19.55	118.722	1242.024	0.410	31.70	49	55	143	147	158	167	223	E
16	J060034-285854	89.7	35.3	202.2	06:00:34.25	-28:58:54.78	1380.526	19.92	24.09	197.870	582.788	0.173	13.35	88	92	33	38	195	210	234	
17	J055851-263952	95.5	70.0	222.2	05:58:51.96	-26:39:52.50	1645.390	7.98	4.05	65.957	1.137	0.063	4.86	95	96	70	70	220	225	10	
18	J061707-272356	34.7	58.3	227.8	06:17:07.75	-27:23:56.08	1718.265	16.62	19.53	303.400	108.781	0.093	7.17	33	36	57	61	215	238	100	
19	J061523-263345	40.1	71.0	232.2	06:15:23.56	-26:33:46.00	1777.077	12.33	11.69	39.574	11.080	0.068	5.28	39	41	70	72	231	234	17	
20	J055847-252521	95.8	88.6	232.3	05:58:47.87	-25:25:21.21	1778.689	23.91	16.14	237.444	182.857	0.115	8.91	93	98	87	90	221	239	140	
21	J060053-214231	88.8	144.3	232.9	06:00:53.96	-21:42:31.77	1785.591	27.95	24.13	211.061	966.685	0.166	12.78	86	92	142	147	223	239	297	E
22	J060444-260638	75.8	78.3	233.2	06:04:44.15	-26:06:38.94	1789.930	20.11	19.90	224.253	410.121	0.155	11.93	74	78	76	80	224	241	245	
23	J060107-233957	88.0	115.0	235.5	06:01:07.39	-23:39:57.83	1820.476	31.94	32.09	277.018	3195.852	0.297	22.92	85	92	112	119	225	246	554	
24	J061538-223529	37.8	130.5	254.5	06:15:38.92	-22:35:29.08	2070.321	12.20	7.70	65.957	5.855	0.070	5.38	37	39	130	131	251	256	17	
25	J061736-230442	31.2	123.0	257.6	06:17:36.87	-23:04:42.45	2112.296	12.13	3.64	39.574	1.194	0.056	4.29	30	32	123	123	256	259	6	
26	J061210-214906	49.6	142.3	270.1	06:12:10.23	-21:49:06.41	2276.118	20.28	15.64	408.931	233.453	0.101	7.78	47	51	141	144	256	287	203	
27	J061719-251534	33.1	90.3	270.4	06:17:19.49	-25:15:34.52	2281.139	12.36	11.64	26.383	8.072	0.075	5.80	32	34	89	91	269	271	14	
28	J060923-223127	59.4	131.9	295.5	06:09:23.70	-22:31:27.05	2612.061	20.60	47.80	52.765	109.813	0.177	13.68	57	61	127	138	293	297	36	
29	J061123-285048	54.1	37.0	296.3	06:11:23.82	-28:50:48.52	2622.770	16.97	47.77	52.765	40.160	0.103	7.92	53	56	30	41	293	297	23	
30	J061919-262155	26.9	73.6	297.6	06:19:19.21	-26:21:55.72	2639.640	12.82	23.59	39.574	10.151	0.091	7.03	26	28	72	77	296	299	14	E
31	J061619-213308	35.1	146.0	298.9	06:16:19.11	-21:33:08.35	2656.418	20.22	7.47	237.444	30.745	0.127	9.81	33	37	145	146	293	311	48	E
32	J055506-295621	107.4	20.8	367.5	05:55:06.56	-29:56:21.28	3561.794	15.70	20.26	39.574	87.311	0.169	13.06	106	109	19	23	366	369	44	
33	J060245-215644	82.3	140.8	434.0	06:02:45.93	-21:56:44.75	4438.143	12.02	12.02	250.635	10.704	0.294	22.73	82	84	139	141	423	442	17	N
34	J055154-220728	120.0	137.8	434.2	05:51:54.49	-22:07:28.09	4441.437	11.28	44.32	79.148	34.948	0.181	14.01	119	121	132	142	431	437	22	
35	J061600-264722	38.1	67.5	547.0	06:16:00.51	-26:47:22.51	5929.367	12.24	7.67	39.574	3.270	0.064	4.95	37	39	67	68	546	549	9	
36	J055212-291700	117.0	30.5	727.0	05:52:12.81	-29:17:00.56	8303.630	3.67	16.10	342.974	142.796	0.479	36.98	117	117	29	32	714	740	102	

Figure 4: Typical output to screen at completion of a *Duchamp* run, showing the range of parameters calculated. Other possible outputs include formats for use with Karma software or Virtual Observatory applications.

Obtaining *Duchamp* and Further Development

Duchamp is easy to install and run, and is available from **http://www.atnf.csiro.au/people/Matthew.Whiting/Duchamp** Any and all feedback is welcome!

Duchamp will continue to be developed, particularly in response to users' requirements. It will also form the basis for analysis software under development for use with the Australian SKA Pathfinder.

CSIRO

References

[1] Miller et al 2001, AJ 122, 3492.

[2] Using the technique of Lutz 1980, The Computer Journal 23, 262

[3] Starck et al 1997, ApJ 482, 1011 is a good example of the technique.

Duchamp makes use of the PGPLOT, CFITSIO and WCSLIB libraries. For more information, see the User's Guide, or the forthcoming paper.

Acknowledgements

The name *Duchamp* comes from the renowned Cubist & Dadaist artist Marcel Duchamp, who pioneered the art of the readymades, or 'found objects'.

The logo at top right is appropriated from Duchamp's 1913 readymade *'Bicycle Wheel'*.

Further information

contact: Matthew Whiting
phone: +61 2 9372 4683
email: Matthew.Whiting@csiro.au
web: www.atnf.csiro.au/people/Matthew.Whiting/Duchamp

www.csiro.au