FHT 2004

Tabular Document Image Registration Using the Fourier-Mellin Transform

Luke Hutchison (lukeh@byu.net) Computer Science, BYU Provo

Abstract. *Image registration*, or the detection and reversal of the transformation that best maps one image to another, is a useful tool in digital processing of microfilm images. Having properly-aligned microfilm images can help in manual and automated content extraction, document zoning, and compression of batches of related tabular document images. An image registration algorithm is presented for quickly identifying the global affine transformation (rotation, scale, translation and/or shear) that maps the tabular template in one document to another, using the Fourier-Mellin Transform (from signal processing theory). This registration method deals with all components of the affine transform in a uniform way, by recovering each component in the frequency domain. This contrasts with current techniques for document image registration, which usually rely on a mix of unrelated techniques for different transform components. Each component of the affine transform is made completely separable from the others, dramatically reducing the total parameter space of the problem. Registration is limited to foreground pixels (the document form and printed text) through the introduction of a novel, locally adaptive foreground-background segmentation algorithm, based on the median filter. Common problems with Fourier-Mellin image registration are eliminated by exploiting the properties of tabular document images. An original method is also presented for automatically

the properties of tabular document images. An original method is also presented for automatically obtaining blank form templates from a set of registered document images, by computing the median value across all registered documents at each pixel. Finally, image registration is demonstrated to be an effective tool for predictive compression of batches of document images which share the same template. The presented registration algorithm is able to handle a wider range of transformation types than most document image registration systems, which usually only perform deskewing, and is shown to be reliable and robust in registering several image datasets.

to be remadie and robust in registering several image datasets.

-	The undermentioned Houses are situate within the Boundaries of the												[Phys. 2
	Civil Parish [or Township] of Municipal				h of Municipal Ward of	100 m	Parliamentary Borough		Tow	n or Village or Hamlet of	Urban Sanitary District of	Rural Sanitary District of	Ecclesiastical Parish or District of
	No. of Schedule	BOAD, STREET, &c., and No. or NAME of HOUSE	USE HOUSE		NAME and Surname of each Person	RELATION to Head of Family	CON- DITION as to Marriage		E last thday of	Rank, Profe	ession, or OCCUPATION	WHERE BORN	(1) Deaf-and-Dum (2) Blind (3) Imbecile or Idi (4) Lunatio
ı	that	describer 11 de	1	(B)	Mr ou so distill a seal	- 100	marriage	Males	Females	1800000	Maria Laboratoria de la Companya de	#2 min 10 10 11	(4) Lunatio
1	PAR .	The medicinal set	1	-	Maria de la companya del companya de la companya de la companya del companya de la companya de l	No. of the second	Marie	1000 No.3	Lie	10000 Bernell 10	Carlo Contract Contract	I Bellines - Wanterwill Biris Buri - Lis - 1880	11-1111 -
	112	Million of the second	0	-	1000 10 de porto 1000 10 de mo-	- Maria	Jane .	27.90 Vi.1	200	THE PARK R	Aller of the second	28-2357-212111	Velia -
ı	1	58046 2012 (September 1997)	1		Bed alicenter	- Bid wood	7	110	1900	TO SELECTION OF THE SEL	VARIANCE SEE	Property St.	Children Comment
ı	67	100165 455	1		All and second	St.	1/4/1	100	100	2000 and 200	State State State	19.20 To 12.11	26.60
ı	18/1	interest of the	1	-	Commence of the Street	1 - All Comments	200	16	184	The same	Alexander Company	2887-116-14	archi.
1	1000	The same	1		William Hard all Andrews	2000	996	N/S	116	A STATE		The state of the s	College Co.
ı	1/1	3/2	1		Mandalarte 12, 18 h	Mary	196	1/3	110	THE DOLLARS	But Brown Land	Distant Company	Mills and
	200	1964 120 - 110	1	1	All Shirter	6	1/2	160	111	03/03-03/03/03/	Will And Stylowers		
	11/10	76.6	1	12	But Soll Walle	1300	780	25	166	300 S. S. S. S.	Strange In X	FD3104532224	addition to the same
	116	12 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	1	1		- Maria	786	10	62	1 Jan 2020		Experience Contraction	44/20011
	110	The state	1	12	May 12 2019 11 2 21	Allert de	70%	16	144	192 Maria 544	21.5 Market X	110000 120 11 11 11 11 11	list of
	1919	(A2 15-15-12 - 17	1		35,43,4,4,4,4,4,4,4,4,4,4,4,4,4,4,4,4,4,	- Miles	Altra .	20	166	Mas Vais	Washington and the	10-22-5-12-5-12-5-5-5-5-5-5-5-5-5-5-5-5-5-	dele
	150	FAMILION STATES	1		Marie Salah Salah	11 11 11	196	Au	100		CV with the said	Burney Daniel	thing -
	field.	Marinet Bridge !	0		Millian Show	Marie	76	200	10	FRED LES	and the state of t	ASSELLATION AND A	allo - 7
	BA	William Samuel	1		Miller and Market State Service	March	360	No.	1675	Philasipa	Market Market Co.	Market State State	United States of States of States
	155	1800 11 -1-	11.	12	Mary Mary Com	Marie	984	362	84	0/41/1	the said the	HONET BULLEY	alfan De
	3,48	1610 1	1		Maria Maria	is Marie	Maria	38.5	160	THE SERVE	Making Water Control	Printer 12 10	and a
	960	Wall and the same	13		Madeline	i Maril	100	1/2	260	W. 182 182	Manual Manual	All Dietar Harris	district the same of the same
	1165	BB B B B B B B B B B B B B B B B B B B	1	100	Mary State Jane	de Mario	Maria	16	14.7	America	White the work of	Matheway Long La	toke - CA
	1800	1371 1421 145	111	12	Para Maria	1 Same	200	30	20	Aller dille	English State State St.	Market Selled Con	the same
١	1/10	Will down a market	1		March Hilliams	a March	1846	19	166	Sall Ballet	Salatia State Sept	Battern William	white the same of
1	898	111 10121	1	2	Browning To be had	March	28	10	64	Addition to	Marine de la como	Harrist March 19	affect the state of
	6311	Man Barthart 2 .	1	1 2	Musik Barre	- Marie	1880	19	13%	170 500	Market D. William	Harris Harris Harris	diff.
ı	9/			1	Mary Mary		1300	000	1			A Charles and the	
	15	Total of Houses.	1	1/10	Tot	tal of Males and	Females	14	14			dille.	