BLIND DATE: USING PROPER MOTIONS TO DETERMINE THE AGES OF HISTORICAL IMAGES

Jonathan T. Barron[1,2], David W. Hogg[3], Dustin Lang[1], and Sam Roweis[1,4]

[1] Department of Computer Science, University of Toronto, 6 King's College Road, Toronto, Ontario M5S 3G4, Canada
[2] Department of Computer Science, The Courant Institute of Mathematical Sciences, New York University, 715 Broadway, NY 10003, USA
[3] Center for Cosmology and Particle Physics, Department of Physics, New York University, 4 Washington Place, NY 10003, USA; david.hogg@nyu.edu
[4] Google Inc, Mountain View, CA, USA

ABSTRACT

Astrometric calibration is based on patterns of cataloged stars and therefore effectively assumes a particular epoch, which can be substantially incorrect for historical images. With the known proper motions of stars we can "run back the clock" to an approximation of the night sky in any given year, and in principle the year that best fits stellar patterns in any given image is an estimate of the year in which that image was taken. In this paper, we use 47 scanned photographic images of M44 spanning the years 1910–1975 to demonstrate this technique. We use only the pixel information in each image; we use no prior information or metadata about image pointing, scale, orientation, or date. Blind Date returns date metadata for each input image. It also improves the astrometric calibration of the image because the final astrometric calibration is performed at the appropriate epoch. The accuracy and reliability of Blind Date are functions of image size, pointing, angular resolution, and depth; performance is related to the sum of proper-motion signal-to-noise ratios for catalog stars measured in the input image. All of the science-quality images and 90% of the low-quality images in our sample of photographic plate images of M44 have their dates reliably determined to within a decade, while 75% of our science-quality images have their dates reliably determined to within 3 years. This method has the potential to insert "lost" data into incomplete sets of time-domain imaging, improving constraints from historical data on transients, variable stars, stellar motions, and minor planet orbits.

Key words: astrometry – catalogs – methods: statistical – stars: kinematics – techniques: image processing – time

1. INTRODUCTION

Astronomy needs well-calibrated data to make precise measurements, but also wants to make use of large data sources that are poorly calibrated. Unreliable data sets such as historical archives, amateurs' collections, and engineering data contain important information, especially in the time domain. Astronomy needs methods by which data of unknown provenance, quality, and calibration can be vetted, calibrated, and made reliably useable by the community.

There is an enormous amount of information about proper motions, stellar and active galactic nucleus (AGN) variability, transients, and solar system bodies in historical plate archives and the collections of good amateur astronomers. The Harvard College Observatory Astronomical Plate Stacks[5] alone contain enough photographic exposures to cover the entire sky 500 times over, and span many decades with good coverage and imaging depth. However, in many cases, it is challenging to use these images quantitatively. Often the details of observing date, telescope pointing, bandpass, and exposure time are lost because the logs have been lost, because the information was written incorrectly or illegibly, or because it was difficult or expensive to associate each image with the appropriate record in the log.

The astronomical world is moving toward the development of a virtual observatory (VO), in which a heterogeneous set of data providers communicate with researchers and the public through open data-sharing protocols.[6] These protocols can easily be spoofed—intentionally and unintentionally—and permit the dissemination of badly calibrated, erroneous, or untrustworthy data.[7] Indeed, the lack of a "trust" model may compromise the VO's goals of making it easy for astronomers to discover and use wide varieties of input data; without trust, all the time that is saved in searching for data has to be spent in verification and tracking of provenance.

Amateur astronomers and educators at college and planetarium observatories are, in many cases, well equipped and can take science-quality observations, especially for time-domain science. In addition, most of these astronomers are interested in contributing to research astronomy. However, it is challenging at present for these potential data providers to provide their data to the community in a form that is trivially useable by the research community. These observers need automated systems to calibrate their data and hardware, and to package the data and metadata in standards-compliant forms. Even professional observatories often produce data with incorrect or standards-violating metadata because of telescope faults, software bugs, or the youth of most standards and conventions for digital data formats.

We have begun a large project (*Astrometry.net*) to vet, restore, determine, and package in a standards-compliant form the calibration metadata for astronomical images for which such information is lost, damaged, or unreliable (D. Lang et al. 2008, in preparation). Our system can astrometrically calibrate an image (i.e., determine the world coordinate system) by using the information in the image pixels alone. It starts by identifying asterisms that determine astrometric calibration. Once the astrometry is correct, the sources in the image can be identified with catalogs and other calibration metadata can be inferred through quantitative image analysis. In the process of testing and running this calibration system, we have indeed

---

[5] http://tdc-www.harvard.edu/plates/
[6] http://www.ivoa.net/

[7] http://www.ivoa.net/Documents/Notes/IVOArch/IVOArch-20040615.html

confirmed that many historical and amateur images—and even some scientific images from modern professional facilities—have missing or incorrect astrometric metadata; automated calibration, as a vetting step at the very least, is essential for all data sources.

For the same reason that the time domain is interesting, it can also be used to calibrate the *date* that an image was taken. Stars move and vary, so the particular configuration and relative brightnesses of the stars in an image provide, in principle, a measure of the time at which the image was taken. Here we show that stellar motions cataloged at the present day can be used to age-date plates from a plate archive to within a few years, even for very old plates. This new capability takes the *Astrometry.net* project a small step closer to being a comprehensive image metadata vetting and automated calibration system.

2. INPUT DATA

The larger goal of our project is to add calibration metadata to data of unknown provenance. For this reason, the system begins truly "blind" in the sense that we ignore all metadata associated with each input image, and start with only the image pixels themselves. We calibrate the images using the *Astrometry.net* astrometric system; this calibration provides output that is taken as input data for the Blind Date analysis.

For each automatically detected star i in the image, there is a centroid (x_i, y_i) in the input image, measured in pixels. For each source, this centroid is the location of the maximum of a second-order polynomial (generalized parabolic) surface fit to the area immediately surrounding the center of the image star. The fit also provides uncertainties $(\sigma_{xi}, \sigma_{yi})$ in these centroids. For unsaturated stars these are taken (arbitrarily) to be 1 pixel, and for saturated stars (which are common in the digitized plate data) these are taken to be one-third of the radius of the saturated region. These are overestimates, since the stars are detected at good signal-to-noise ratio (S/N); furthermore, in the case of saturated stars, it is possible for the saturated "disk" to be nonconcentric with the true centroid.

The system also provides a rough world coordinate system (WCS) for the image; that is, it provides first guesses at two functions: $\text{WCS}_{xy \to \text{RD}}(x, y)$, which transforms points from the image plane in pixels into celestial coordinates in angular units, and the inverse, $\text{WCS}_{\text{RD} \to xy}(\text{R.A., Decl.})$. The pixel scale θ_{pix} (which we will use below), measured in angle per pixel, is derived from these functions. Strictly, θ_{pix} is a function of the position in the image, but in typical science images it does not vary substantially.

The WCS effectively identifies the sources from the USNO-B Catalog (Monet et al. 2003) that are in or likely to be in the image. For each catalog "star" j (the USNO-B Catalog contains both stars and compact galaxies) inside the image, the catalog contains a J2000 celestial position $(\text{RA}_j, \text{Dec}_j)$ on the celestial sphere (extrapolated to epoch 2000.0) measured in angular units, an uncertainty $(\sigma_{\text{RA}j}, \sigma_{\text{Dec}j})$ in that position, a proper motion $(\mu_{\text{RA}j}, \mu_{\text{Dec}j})$ measured in angle per time, and an uncertainty $(\sigma_{\mu\text{RA}j}, \sigma_{\mu\text{Dec}j})$ in that proper motion.

The USNO-B Catalog uncertainties required some processing and adjustment. For the sake of clarity and simplicity, we make as few assumptions as possible in transforming the uncertainties. Our approach here is not intended to be definitive. Many of these entries in the catalog have values of zero for the uncertainty of position or proper motion. In the case of zero-valued uncertainty in position, we assume that the uncertainty is below the precision of 0.002 arcsec at which the catalog was

reported. We therefore set all zero-valued position uncertainties to one-half of the precision (0.001 arcsec). For entries with a zero-valued uncertainty in proper motion, there is more we need to consider; a nonzero-valued proper motion paired with a zero-valued uncertainty indicates that the uncertainty is below the precision of the catalog, and therefore should be set to half of the precision. A zero-valued proper motion paired with a zero-valued uncertainty indicates that the proper motion of the entry could not be measured accurately (D. Monet 2008, private communication). We therefore set the uncertainty in the proper motion for such entries to three times the median value of nonzero proper motion uncertainties. This captures the idea that, generally speaking, we are significantly more uncertain about the proper motion of such entries than we are about most other entries. A more principled approach could certainly be attempted, but we found that this worked well enough for our purposes.

Much work has already been done by the *Astrometry.net* team to identify spurious sources in the USNO-B Catalog that appear to have been created by diffraction spikes and reflection halos (Barron et al. 2008). For the purposes of this project, we ignore the entries in the catalog which have been flagged as spurious. Testing has shown that ignoring these sources generally improves the accuracy of our results.

Using a technique that will be described in a future paper from the *Astrometry.net* team, we are able to estimate the bandpass of each image being processed, in that we determine which bandpass of the USNO-B Catalog most closely predicts the brightness ordering of the stars in the image. Though this technique is still in an experimental stage, its results are not very controversial; all of Harvard's images of M44 appear to best match the blue bands (O and J emulsions) of the USNO-B Catalog. This finding is reinforced through experimentation with the manual setting of each image's band: on average, Blind Date performs better on these images using the J emulsion of the catalog than on any other band.

Assuming that we have estimated the bandpass correctly, the N stars in the image should correspond—roughly—to the N brightest catalog stars that lie within the area of the image. For this reason, we only use the N brightest catalog stars in what follows.

3. METHOD

We use the USNO-B Catalog positions and proper motions to "wind" the N catalog stars backwards and forward through time along the celestial sphere. Once the catalog has been adjusted, we can use the image WCS to project the catalog entries onto the image plane, making a synthetic catalog for that image at that time, in image coordinates. We then attempt to fit the image stars to that synthetic image of the moved catalog stars. In this fit, we permit polynomial warping of the image relative to the catalog in the image plane. We specify a scalar objective function that is minimized when the image is "best" warped. We warp the input image to the catalog "wound" to different times, and use the time at which the best-fit objective function value is smallest as the estimate of the time that the image was taken.

3.1. Winding Back the Catalog

We estimate the celestial coordinates of the catalog star j at the arbitrary date t, and then project them onto the image plane:

$$(u_j, v_j) = \text{WCS}_{\text{RD} \to xy}(\text{RA}_j - \hat{\mu}_{\text{RA}j}[t - (2000.0\,\text{yr})],$$
$$\text{Dec}_j - \mu_{\text{Dec}j}[t - (2000.0\,\text{yr})]), \tag{1}$$

where we have adjusted the $\mu_{\mathrm{RA}j}$ proper motion vector components and their associated uncertainties into coordinate derivatives $\hat{\mu}_{\mathrm{RA}j}$ by

$$\hat{\mu}_{\mathrm{RA}j} = \cos(\mathrm{Dec}_j)\,\mu_{\mathrm{RA}j} \qquad (2)$$
$$\hat{\sigma}_{\mu\mathrm{RA}j} = \cos(\mathrm{Dec}_j)\,\sigma_{\mu\mathrm{RA}j}.$$

We estimate the uncertainty of the location on the image plane of the catalog star j at year t with

$$\sigma_{uj} = \frac{1}{\theta_{\mathrm{pix}}}\sqrt{\sigma_{\mathrm{RA}j}^2 + (\max(|t - t_{\mathrm{early}}|, |t - t_{\mathrm{late}}|)\,\hat{\sigma}_{\mu\mathrm{RA}j})^2}$$

$$\sigma_{vj} = \frac{1}{\theta_{\mathrm{pix}}}\sqrt{\sigma_{\mathrm{Dec}j}^2 + (\max(|t - t_{\mathrm{early}}|, |t - t_{\mathrm{late}}|)\,\sigma_{\mu\mathrm{Dec}j})^2},$$
$$(3)$$

where t_{early} and t_{late} are (roughly) the observation dates for the images from which the USNO-B Catalog was constructed, which, in this patch of the sky, are 1955.0 and 1990.0, respectively. All positions in the catalog, however, were extrapolated to the year 2000.0 (epoch and equinox). Although we are given each catalog star's location at the year 2000.0, we know that each star's measured location is, in general, more accurate between the two epochs in which the images were taken and less accurate at years further from that range. Therefore, when "winding" locations through time, we look at the difference between t and the year 2000.0; when "winding" uncertainties through time, we look at the maximum distance between t and both t_{early} and t_{late}. This is equivalent to using $(t_{\mathrm{early}} + t_{\mathrm{late}})/2$ as our reference year and assuming a nonzero uncertainty on the measurement of each star's proper motion at that reference year.

Note that in our notation for (u_j, v_j) and $(\sigma_{uj}, \sigma_{vj})$, we do not reference t. This is because once the catalog has been wound through time and projected onto the image plane, we consider the time to be fixed. Note that in Sections 3.2 and 3.3, the time will remain fixed, and therefore t is not mentioned in any of the notations except for in $\chi^2(t)$.

3.2. Objective Function

Determination of the image coordinate system and date involves finding parameters—astrometric parameters and the date—that optimize an objective function. The choice of the objective function is therefore the fundamental scientific decision in the project.

We seek an objective function that has the following properties, listed in a rough order of priority. The function must decrease as image-coordinate distances between the catalog and image stars decrease. The function must be insensitive to anomalous outliers and more sensitive to well-measured stars than to poorly-measured stars. The function must be some approximation to a likelihood function or have some equivalent justification so that changes in the function with respect to parameters can be interpreted in terms of uncertainties in those parameters. The function ought to be differentiable and second-differentiable with respect to all fit parameters (in particular, time and astrometric calibration). The function should be easily optimized. We have identified an objective function that has all of these properties; it is so similar to the least-square function that we call it a "modified chi-squared" and denote it as "χ^2."

For all $i = 1{:}N$ and $j = 1{:}N$ we compute d_{ij}, the Euclidean distance between the image star i and catalog stars j in the image plane:

$$d_{ij} = \sqrt{(u_j - x_i)^2 + (v_j - y_i)^2}. \qquad (4)$$

We also estimate the uncertainty $\bar{\sigma}_{ij}$ of each pair's distance measurement, using the previously-defined values for $(\sigma_{xi}, \sigma_{yi})$ and $(\sigma_{uj}, \sigma_{vj})$. We calculate the combined uncertainty for the pair in x and y at time t, and then simply take the mean as an estimation of the uncertainty of that pair:

$$\bar{\sigma}_{ij} = \tfrac{1}{2}\left(\sqrt{\sigma_{xi}^2 + \sigma_{uj}^2} + \sqrt{\sigma_{yi}^2 + \sigma_{vj}^2}\right). \qquad (5)$$

This is not a rigorous estimate of the uncertainty; it is just a heuristic that accounts for contributions to the uncertainty from both the image and the catalog.

We define a weighting function ρ for each pair that returns a value between 0 and 1 based on the ratio of the distance of a pair to the uncertainty of that pair:

$$\rho(d, \bar{\sigma}) = \frac{1}{1 + \left(\frac{d}{\bar{\sigma}}\right)^2} \qquad (6)$$

(see also Figure 1). This function has the property that outliers are down-weighed in a smooth manner; it causes the influence of an image–catalog pair to smoothly drop to zero at large displacement. This permits us to avoid the discontinuous optimization problem of sigma clipping, which is the standard "robust estimation" technique in common use in astronomy applications. The weighting function ρ has a number of properties which make it well suited to our purpose:

$$\rho(0, \bar{\sigma}) = 1 \qquad (7)$$

$$\left.\frac{d\rho}{dd}(d, \bar{\sigma})\right|_{d=0} = 0 \qquad (8)$$

$$\lim_{(d/\bar{\sigma}) \to \infty} \rho(d, \bar{\sigma})\left(\frac{d}{\bar{\sigma}}\right)^2 = 1. \qquad (9)$$

We use this weighting function to assign a weight w_{ij} to all pairs as follows:

$$w_{ij} = \rho(d_{ij}, \bar{\sigma}_{ij}). \qquad (10)$$

Our final objective function is

$$\chi^2(t) = \sum_{ij} w_{ij}\left(\frac{d_{ij}}{\bar{\sigma}_{ij}}\right)^2, \qquad (11)$$

where the sum is over all possible image–catalog pairs.

For "good" (small-separation) image–catalog pairs, the weight function is near unity (Equation (7)) and has near-zero derivatives (Equation (8)), so small changes in separation do not enter strongly into derivatives of the objective function. For "bad" (large-separation) image–catalog pairs, the pair's contribution to $\chi^2(t)$ is nearly constant (Equation (9)). This makes optimization and interpretation of our objective function very similar to optimization and interpretation of a chi-squared fitting system. This weighted $\chi^2(t)$ objective function could also be interpreted as a Geman–McLure error function. In this framework, optimizing the objective function is equivalent to robust M-estimation (Hampel et al. 1986).

In constructing the modified chi-squared, we use—in principle—all image–catalog pairs, irrespective of their separation in image coordinates. The fact that the contribution of a pair to the objective function quickly converges to 1 as the pair becomes highly separated allows us, in practice, to ignore all highly separated pairs. We therefore choose to approximate the contributions of all pairs where $d > 8\bar{\sigma}$ as $\rho(8\bar{\sigma}, \bar{\sigma})\left(\frac{8\bar{\sigma}}{\bar{\sigma}}\right)^2$, which is 64/65. This dramatically speeds up our computation.
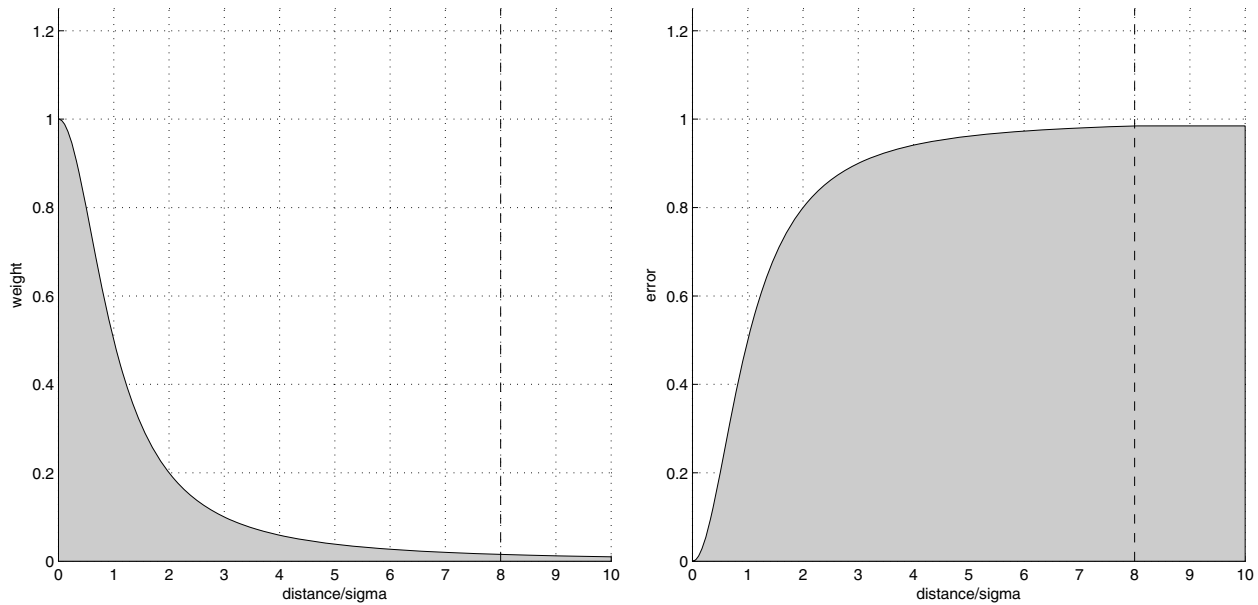
Figure 1. Left: the weighting function vs. a pair's ratio of distance to uncertainty. Right: the corresponding weighted contribution of a pair to the modified chi-squared. The dotted line indicates the point at which we assume that the weighted contribution stops changing.

3.3. Fitting the Image

At this point, we have our image stars on the image plane, our catalog stars (wound to the time of interest) projected onto the image plane, and an objective function that we wish to minimize. We need to find the transformation that we can apply to the image that minimizes the objective function, which we take to be the transformation that best "fits" the image to the catalog.

Testing has suggested that the initial location of the image returned by the *Astrometry.net* solver is close enough to the optimal location that locally minimizing the objective function is sufficient for finding the global minimum, and that we generally do not run the risk of falling into a false local minimum. Therefore, we only present our method for locally minimizing the optimal function through iteratively reweighted least-squares (IRLS). We experimented with techniques such as RANSAC to fit images to the catalog in the face of extreme noise, but no technique was more effective and robust than our IRLS method.

Let us first construct a solution to a simplified version of this problem, in which we know which correspondences are true: we assume that the image point i corresponds to the catalog point i for all $i \in \{1, 2, \ldots, N\}$. Assuming that we are interested in solving for an affine transformation (first-order linear transformation plus shift), we need to find the transformation matrix that best satisfies the following equations:

$$\forall_{i \in \{1,2,\ldots,N\}} \begin{bmatrix} m_x & m_y & t_x \\ n_x & n_y & t_y \end{bmatrix} \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} = \begin{bmatrix} u_i \\ v_i \end{bmatrix}. \quad (12)$$

This can be generalized straightforwardly for higher-order transformations.

The transformation that best satisfies these equations can be found by using a standard least-squares solver. We can then use this transformation to warp all of the image points onto the catalog points (and vice versa), thus solving our simplified problem.

Of course, since we do not know which image stars correspond to which catalog stars, we must include equations for all

image–catalog pairs. We are not interested in the solution to this problem, as it would describe a transformation from *every* image star to *every* catalog star. To specify a transformation that satisfies *likely* image–catalog correspondences, we must use our weighting function to make soft assignments regarding correspondences. We therefore use the following equations:

$$\forall_{i \in \{1,2,\ldots,N\}} \forall_{j \in \{1,2,\ldots,N\}} \begin{bmatrix} \frac{\sqrt{w_{ij}}}{\bar{\sigma}_{ij}} & 0 \\ 0 & \frac{\sqrt{w_{ij}}}{\bar{\sigma}_{ij}} \end{bmatrix} \begin{bmatrix} m_x & m_y & t_x \\ n_x & n_y & t_y \end{bmatrix} \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} \frac{\sqrt{w_{ij}}}{\bar{\sigma}_{ij}} & 0 \\ 0 & \frac{\sqrt{w_{ij}}}{\bar{\sigma}_{ij}} \end{bmatrix} \begin{bmatrix} u_j \\ v_j \end{bmatrix}. \quad (13)$$

We begin our solution by constructing a linear system, which contains all of the previously described (unweighted) equation:

$$\begin{bmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & x_1 & y_1 & 1 \\ x_1 & y_1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & x_1 & y_1 & 1 \\ & & \cdots & & & \\ x_N & y_N & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & x_N & y_N & 1 \\ x_N & y_N & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & x_N & y_N & 1 \end{bmatrix} \begin{bmatrix} m_x \\ m_y \\ t_x \\ n_x \\ n_y \\ t_y \end{bmatrix} = \begin{bmatrix} u_1 \\ v_1 \\ u_2 \\ v_2 \\ \cdots \\ u_{N-1} \\ v_{N-1} \\ u_N \\ v_N \end{bmatrix}. \quad (14)$$

We can write this matrix equation as

$$\mathbf{Ax} = \mathbf{b}. \quad (15)$$

To introduce the weight values described in the equations, we construct our weight matrix $\mathbf{W}$, as follows:

$$\mathbf{W} = \text{diag}\left(\frac{\sqrt{w_{1,1}}}{\bar{\sigma}_{1,1}}, \frac{\sqrt{w_{1,1}}}{\bar{\sigma}_{1,1}}, \frac{\sqrt{w_{1,2}}}{\bar{\sigma}_{1,2}}, \frac{\sqrt{w_{1,2}}}{\bar{\sigma}_{1,2}}, \ldots, \frac{\sqrt{w_{N,N}}}{\bar{\sigma}_{N,N}}, \right.$$
$$\left. \frac{\sqrt{w_{N,N}}}{\bar{\sigma}_{N,N}} \right). \quad (16)$$

Our final matrix equation can then be written as

$$\mathbf{WAx} = \mathbf{Wb}. \quad (17)$$

We then find **x** such that the squared residuals, $(\mathbf{WAx} - \mathbf{Wb})^{\mathbf{T}}$ $(\mathbf{WAx} - \mathbf{Wb})$, are minimized. By construction, **x** describes the transformation that best satisfies all of the equations we previously constructed, and can be found using a standard least-squares solver.

Using the transformation defined by **x**, we can calculate the coordinates of our warped image points:

$$x_i' = m_x x_i + m_y y_i + t_x$$
$$y_i' = n_x x_i + n_y y_i + t_y. \qquad (18)$$

Unlike the simplified version of this problem, the warp found after one iteration is not our final solution. As we warp the image, the distances and weights between the image and catalog points change, and so our objective function is no longer minimized. The solution is obtained by repeatedly recalculating our d_{ij} and w_{ij} values and our weighted least-squares transformation. We perform this iteratively reweighted least-squares operation until the resulting transformations stop changing, which, by construction, is also when our objective function stops changing. The solution that we converge upon is returned as the best fit of the image onto the catalog.

The IRLS method minimizes the objective function, as it was constructed such that the objective function is equal to the sum of the squares of the weighted residuals of the matrix equation:

$$(\mathbf{WAx} - \mathbf{Wb})^{\mathbf{T}}(\mathbf{WAx} - \mathbf{Wb}) = (\mathbf{W}(\mathbf{Ax} - \mathbf{b}))^{\mathbf{T}}(\mathbf{W}(\mathbf{Ax} - \mathbf{b}))$$

$$= \sum_{ij} \left(\frac{\sqrt{w_{ij}}}{\bar{\sigma}_{ij}}(u_j - x_i') \right)^2 + \left(\frac{\sqrt{w_{ij}}}{\bar{\sigma}_{ij}}(v_j - y_i') \right)^2$$

$$= \sum_{ij} \left(\frac{\sqrt{w_{ij}}}{\bar{\sigma}_{ij}} d_{ij}' \right)^2$$

$$= \sum_{ij} w_{ij} \left(\frac{d_{ij}'}{\bar{\sigma}_{ij}} \right)^2$$

$$\approx \chi^2(t). \qquad (19)$$

We say that the residuals only *approximate* the optimal function because we use the weights of the previous transformed image (w_{ij}) and the distances of the next warped transformed image (d_{ij}'). Though this may seem strange, once the IRLS has converged on a solution and the optimal function has stopped changing, the distances in one iteration are effectively equal to the distances in the following iteration.

For numerical stability we always use the initial coordinates when constructing our transformation and only return the final warped image coordinates. The intermediate warped image coordinates are only used for calculating distances and weights. This means that our final image coordinates have only been subjected to one transformation, as opposed to many small transformations.

See Figure 2 for some examples of fit and unfit images against correct and incorrect catalog dates respectively.

3.4. Estimating the Date

Now that we have defined a method for fitting an image to the catalog, and a metric for assessing the degree to which an image can be fit, we can use a number of techniques to estimate the year that gives the best fit. We take that year to be the year in which the image was created.

This problem can be phrased as such: an image has some unknown chi-squared curve $\chi^2(t)$, of which we wish to find the year t_0, the theoretical optimal year, such that

$$t_0 = \arg\min_t \chi^2(t). \qquad (20)$$

In the algorithm we will describe, finding t_0 is not computationally feasible, so we must settle on finding a reasonable approximation of the minimum t^*, such that, given an accuracy threshold δ_χ, we are confident that

$$\|\chi^2(t_0) - \chi^2(t^*)\| \leqslant \delta_\chi. \qquad (21)$$

Once we have found t^* and therefore $\chi^2(t^*)$, we also wish to find the extents of our uncertainty region, (t_-^*, t_+^*), such that

$$t_-^* < t^* < t_+^*$$
$$\chi^2(t_-^*) = \chi^2(t^*) + 1 = \chi^2(t_+^*) \qquad (22)$$
$$\mathop{\forall}_{t \in [t_-^*, t_+^*]} \chi^2(t) \leqslant \chi^2(t^*) + 1.$$

Ideally, we want to find all of these values as accurately as possible while sampling the χ^2 curve as few times as possible. There are a number of methods by which we can accomplish this task, each with different tradeoffs concerning efficiency and assumptions about the shape of the curve.

The simplest method for estimating the origin date is through brute force. We sample our χ^2 curve at regular intervals, and take the year in which our objective function scored the lowest as t^*. We then linearly interpolate along the curve to find the extents of the uncertainty region. This method is terribly inefficient and assumes nothing about the shape of the curve, so we only use it as an approximate ground truth to which we will compare our final algorithm.

Our search algorithm begins with first sampling our χ^2 curve at a very broad, regular interval. Though only one initial sample is required, for the figures shown in this paper we sample the curve at 1900, 1950, and 2000. We take the sampled year with the lowest χ^2 score to be t_n, and we then iteratively refine t_n into t_{n+1} until we believe that we have found t^*.

Our objective function was constructed so that we could efficiently calculate $\frac{d\chi^2}{dt}(t)$ and $\frac{d^2\chi^2}{dt^2}(t)$. The equations for these analytical derivatives are elaborate, so we do not present them here. Since $\chi^2(t^*)$ is a minimum in the χ^2 curve, we can assume that $\frac{d\chi^2}{dt}(t^*) = 0$. We can therefore use Newton's method to find t_{n+1}:

$$t_{n+1} = t_n - \frac{\frac{d\chi^2}{dt}(t_n)}{\frac{d^2\chi^2}{dt^2}(t_n)}. \qquad (23)$$

We iteratively evaluate t_{n+1} until $\|\chi^2(t_{n+1}) - \chi^2(t_n)\| \leqslant \delta_\chi$ holds for two consecutive iterations, at which point we take t_{n+1} as t^*. Since Newton's method generally converges quadratically, this is an extremely fast process.[8]

---

[8] Of course since $\chi^2(t)$ is not globally quadratic, Newton's method is not guaranteed to converge from any starting point. To improve numerical conditioning and ensure that the search remains well behaved in the face of somewhat irregular χ^2 curves, we require that $\frac{d^2\chi^2}{dt^2}(t_n) \geqslant \epsilon$, where $1 + \epsilon$ is the smallest representable number >1 on our machine and we require that $|t_{n+1} - t_n| < 25$ yr. Additionally, if Newton's method appears to be diverging or failing to converge, we find t_{n+1} using a binary-search approach in which we sample the midpoint of the area in which the minimum appears to lie. This collection of restrictions effectively amounts to intelligent gradient descent, which we switch to when Newton's method cannot be performed. This fallback system is rarely required, but does sometimes prevent oscillation and search failure.
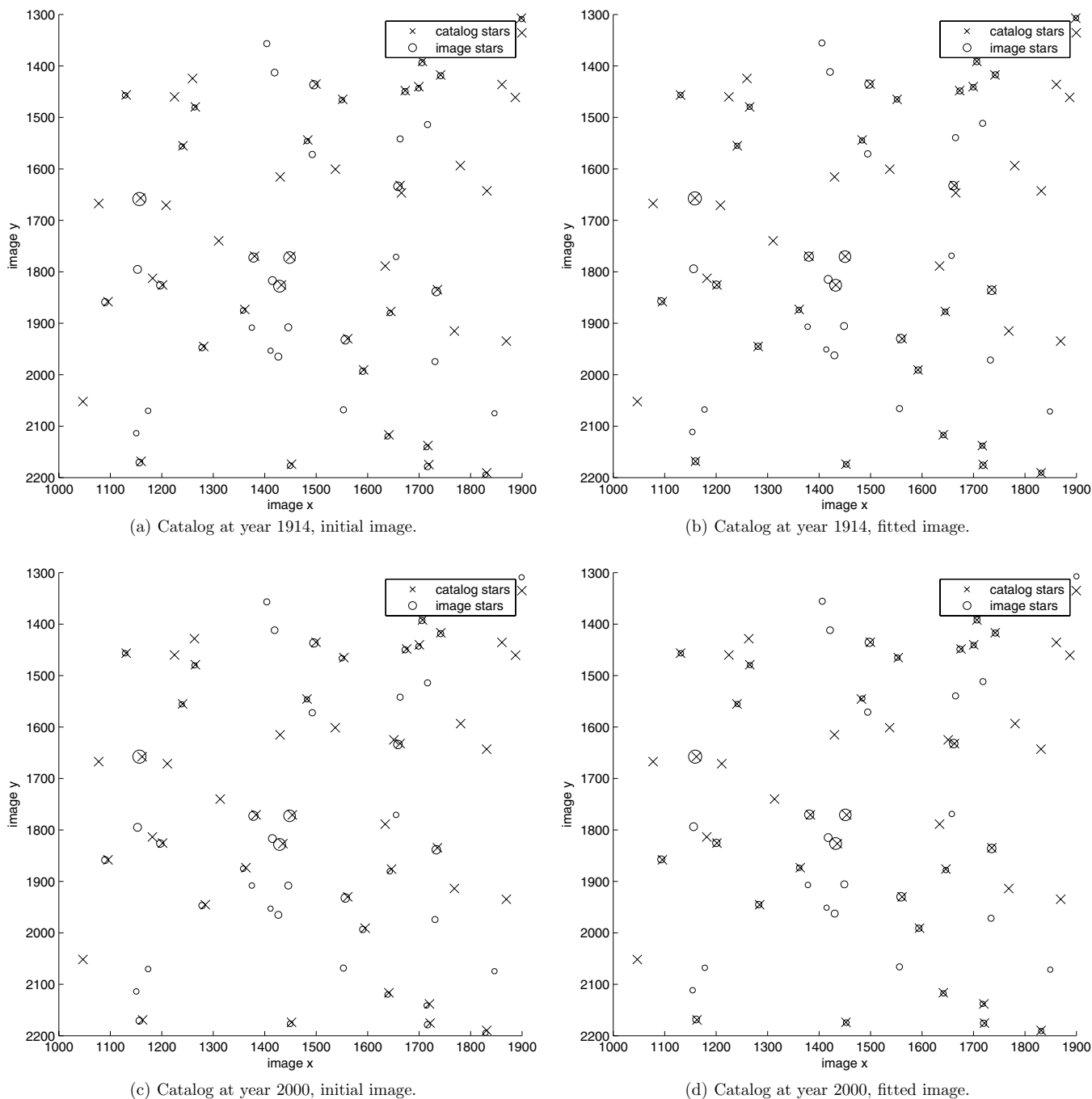
Figure 2. The extracted sources (size is proportional to brightness) from the 900×900 pixel subimage shown in Figure 5. These plots illustrate the error in the initial calibration returned by the *Astrometry.net* solver, as well as the difference in fitting a historical image to the USNO-B Catalog in the year 2000 and to the catalog in 1914, the year which we correctly estimate to be the image's year of origin. (a) Catalog at year 1914, initial image. (b) Catalog at year 1914, fitted image. (c) Catalog at year 2000, initial image. (d) Catalog at year 2000, fitted image.

Once we have found t^*, we can locate t^*_- and t^*_+. We use our modified Newton's method with these new goals:

$$\chi^2(t^*_+) = \chi^2(t^*) + 1$$
$$\chi^2(t^*_-) = \chi^2(t^*) + 1. \qquad (24)$$

This uncertainty region would be the true one-sigma uncertainty region in the limit that the modified chi-squared were the standard linear-fitting chi-squared. Because of the weighting function, in practice this criterion overestimates the one-sigma uncertainty.

We can perform root-finding on these equations using the following formula for iteration:

$$t_{n+1} = t_n - \frac{\chi^2(t_n) - (\chi^2(t^*) + 1)}{\frac{d\chi^2}{dt}(t_n)}. \qquad (25)$$

We begin our two new searches with a sensible initial estimate of the bounds of the uncertainty region, based on our present knowledge of the χ^2 curve. These searches operate under all of the constraints under which the previously detailed search operated, and also converge when $\|\chi^2(t_{n+1}) - \chi^2(t_n)\| \leqslant \delta_\chi$ holds for two consecutive iterations.

Additionally, we can speed up the total search process by using the transformed image points from the previous iteration to find the new transformation for the next iteration. This heuristic
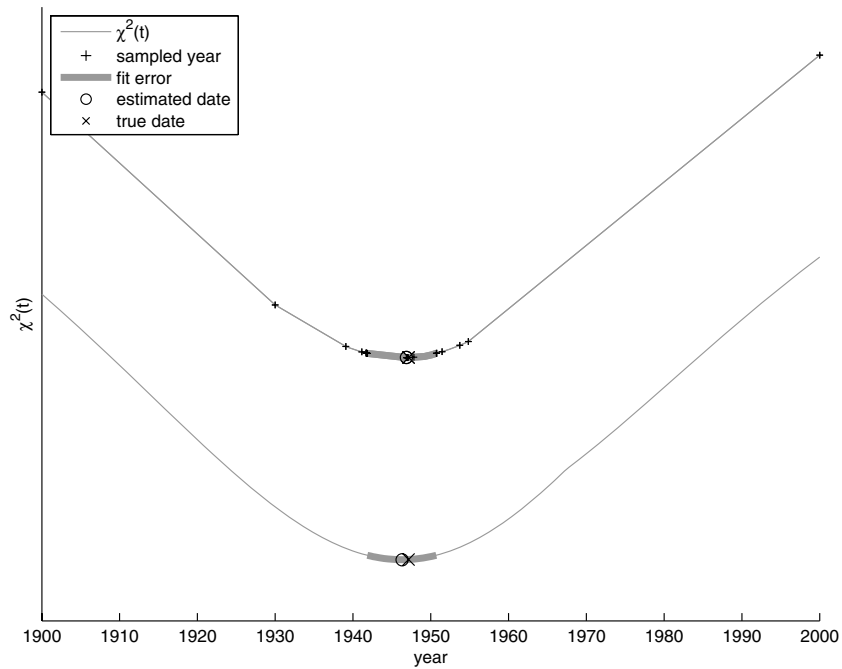
Figure 3. Evaluation of the accuracy of our search algorithm. The top is our modified Newton's method with $\delta_\chi = 10^{-4}$ and the bottom is brute force, with an interval of 0.1 years. The uncertainty regions are identical to within 10^{-3} years, and the estimated dates are within 0.05 years of each other. Brute force took 1001 iterations to achieve this accuracy, while Newton's method took only 23. For the sake of clarity, the curves were vertically separated, and the individual points that were sampled to construct the brute-force curve are not displayed.

means that as the search converges on a final result, the amount of time required to query new years is dramatically reduced. Also, it becomes easier to visualize the search algorithm as a single bidirectional fitting process, in which we repeatedly fit the image to the catalog and the catalog to the image until both fittings have converged. Just as in the previous section, all transformations are constructed using the initial positions of the points, so our final transformation after searching the χ^2 curve is still numerically very stable. Figure 3 shows a comparison of this search algorithm with a brute-force "ground truth." Figure 4 shows the modified χ^2 curves for each image as they were estimated by this search algorithm.

The output of this process is a polynomial description of the image astrometric WCS, a best-fit year value t^*, and an uncertainty region around that value.

3.5. Implementation Notes

Blind Date's two primary performance bottlenecks are calculating the distances between the image and catalog points and solving the weighted least-squares problems. In both cases, we are able to use the properties of our weighting function to construct approximate solutions that very effectively approximate the true solution.

Our algorithm theoretically requires us to repeatedly calculate the distances of all image–catalog pairs. However, due to the properties of the weighting function as described in Section 3.2, we do not need to know the distances of significantly separated pairs. Because the transformation applied at each iteration tends to be very small, we can generally assume that significantly separated pairs in one iteration will also be significantly separated in the next iteration. This allows us to do one initial calculation of all image–catalog distances, but in later iterations only calculate the distances of image–catalog pairs that will likely cause a change in the value of the objective

function. This is a rough heuristic, so we safeguard ourselves by manually recalculating all image–catalog distances every ten iterations, as well as whenever the IRLS begins to converge. This dramatically speeds up our algorithm, while producing results nearly identical to the naïve brute-force approach.

We use the aforementioned properties of our weight function to determine if an image–catalog pair should be considered in the weighted least-squares calculation. A highly separated pair always contributes a nearly-constant value to the residuals, and therefore can be safely ignored. This gives us a slight performance boost.

We require an additional threshold for the difference in the optimal function from one iteration to the next, which determines when our IRLS operation has converged. We use the very conservative value of 10^{-4} as the maximum amount that the χ^2 score of the final IRLS iteration is allowed to change from those of the previous two IRLS iterations.

4. RESULTS

Harvard's interface for accessing its scanned plates of M44 made it difficult to obtain more than 3000×3000 pixel subsets of the images, even though the entire plates are significantly larger. The interface for downloading the images did not provide an obvious mechanism for selecting the same 3000×3000 pixel subsets of each image, which means that such a selection was done by hand and is therefore not very accurate. Many of the plates suffer from the various sources of noise typical of historical imagery: some are multiple-exposures, badly out of focus, badly saturated, or cracked, and some contain handwritten labels, digital scanning artifacts, and bad trailing. For the sake of fairly assessing Blind Date's performance, we split the images into two sets: 27 "science-quality" images and 20 "low-quality" ones (see Figure 5 for examples). For our convenience, we used the JPEG versions of the images,
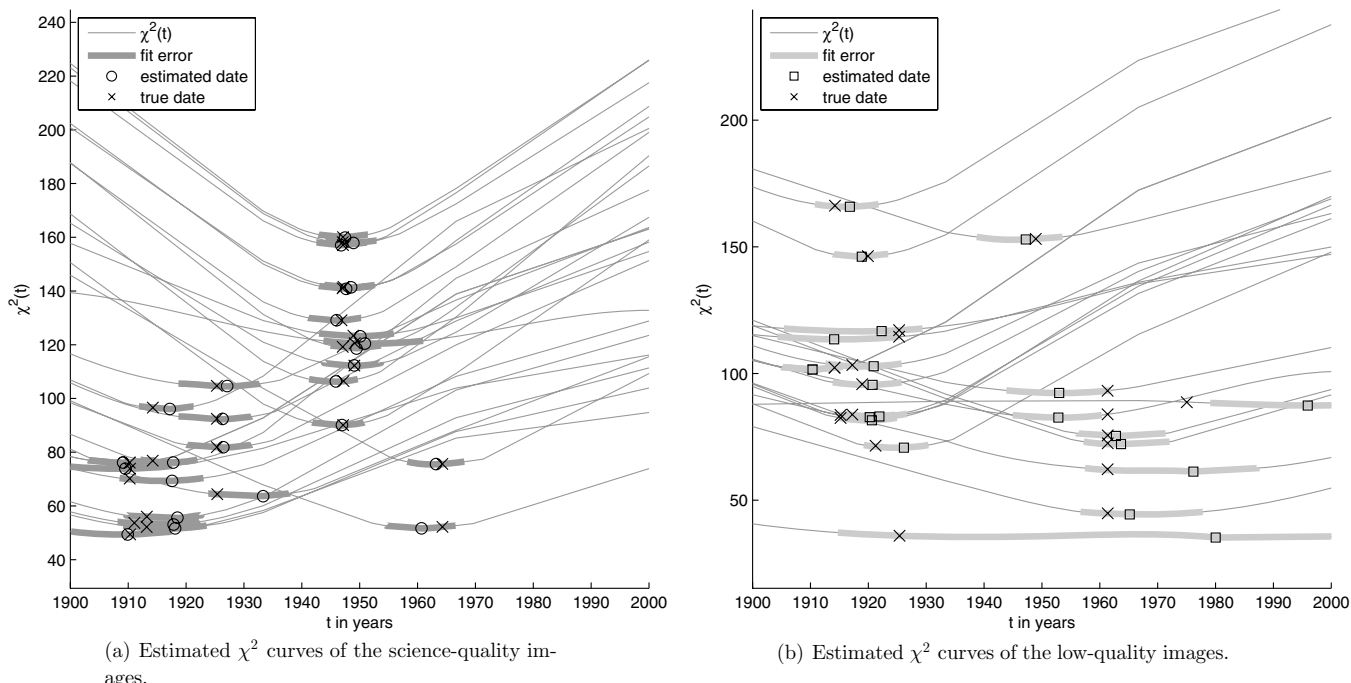
(a) Estimated χ^2 curves of the science-quality images.

(b) Estimated χ^2 curves of the low-quality images.

Figure 4. The estimations of each image's χ^2 curves generated by our modified Newton's method, for both data sets. The vertical axes of these plots do not show the constant contributions of image–catalog pairs whose separations are always too large to be directly calculated, as including those would make the curves excessively vertically separated, rendering these plots incomprehensible. (a) Estimated χ^2 curves of the science-quality images. (b) Estimated χ^2 curves of the low-quality images.
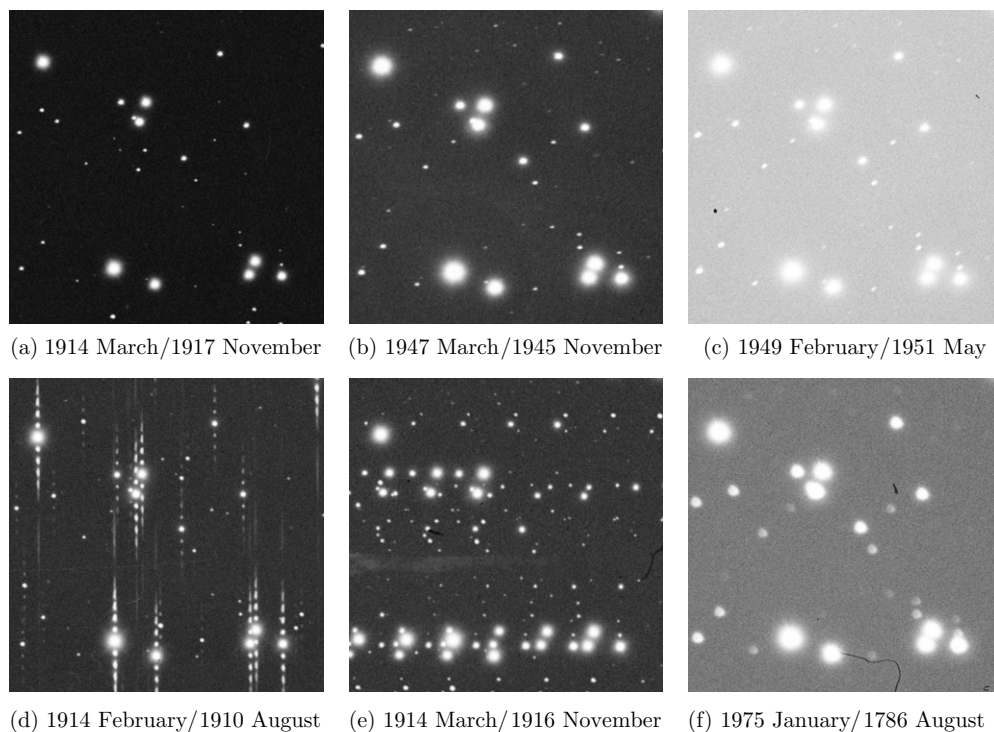

(a) 1914 March/1917 November   (b) 1947 March/1945 November   (c) 1949 February/1951 May

(d) 1914 February/1910 August   (e) 1914 March/1916 November   (f) 1975 January/1786 August

Figure 5. A series of 900×900 pixel subsets of our 3000×3000 pixel images. The top three images are from our set of science-quality images and the bottom three images are from our set of low-quality images. The captions are of the form "true date/estimated date." All but (f) have estimated dates that lie within the uncertainty region. (a) 1914 March/1917 November, (b) 1947 March/1945 November, (c) 1949 February/1951 May, (d) 1914 February/1910 August, (e) 1914 March/1916 November, (f) 1975 January/1786 August.

which probably introduce some minor noise in the form of compression artifacts. Harvard graciously provided ground-truth dates for each image, which were presumably taken from logs or from writing on the plates. We take these dates to be true. The ground-truth dates range from 1910 to 1975, but the dates are not uniformly distributed. See the distribution of

images along the y-axis of Figure 6 for a demonstration of this clustering.

The tests were performed using the modified Newton's method, with affine distortions and an accuracy threshold δ_χ of 10^{-4} year. Tests were done on a 2007 Macbook with a 2 GHz Intel Core 2 Duo processor and 2 GB of RAM. The median

(a) Performance on the science-quality images.
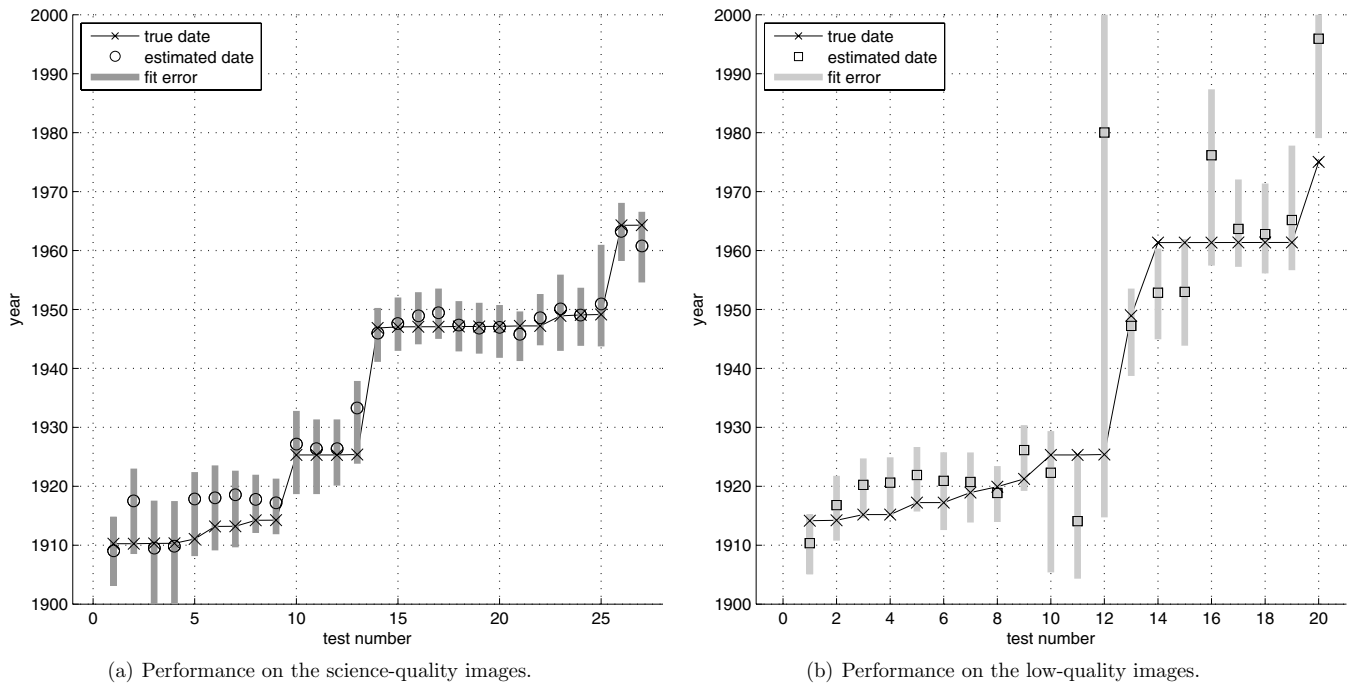

(b) Performance on the low-quality images.

Figure 6. An informative visualization of performance for both data sets. The *x*-axis is an arbitrary index denoting the test image (images were sorted by true year) and the *y*-axis shows each image's true year, estimated year, and uncertainty region. The last image in (b) is not shown, as it is estimated as originating before 1900. (a) Performance on the science-quality images. (b) Performance on the low-quality images.

Table 1
Accuracy of Estimated Dates for the Two Subsets of Data

	Mean Year Error (bias)	Median Absolute Error	Fraction Within Uncertainty
Science-quality	2.01	1.66	27/27
Low-quality	−6.46	4.63	18/20

Note. See Figures 6 and 7 for a more detailed visualization of the results.

runtime for estimating each image's date was ∼2.6 s per image, after source extraction and *Astrometry.net*'s initial calibration. Results are shown in Table 1 and Figure 7.

Though the results shown were generated by fitting an affine (linear) transformation in image coordinates, we experimented by increasing the order of the polynomial warp being fitted. The results were very similar to those obtained using affine transformations, although the median absolute error for science-quality images dropped to 1.01 years for second-order transformation and to 0.99 years for third-order transformations. The median absolute error for low-quality images also decreased slightly as the order was increased, as did the bias for science-quality images.

Additionally, we tested Blind Date on the five USNO-B Catalog source images of M44 that we were able to retrieve from the US Naval Observatory Precision Measuring Machine Data Archive, and on the Sloan Digital Sky Survey (SDSS) (York et al. 2000) image of M44. The dates of the USNO-B Catalog source imagery were estimated accurately (all within 6 years of the true dates and all within the uncertainties), which is as good as we would expect the performance to be, given the relatively low resolution (3.2 arcsec pix^{-1}) of the source imagery that we were able to obtain. The SDSS image was estimated to have been taken in late 2004, but was actually taken in 2006.

We assembled an alternate test-bed of ten amateur images of M44 that we were able to find online. The Web sites on which we found most of our imagery did not explicitly note the date that the image was taken, so we were forced to use the "date" tag in each image's EXIF metadata as the ground truth. For many of the images, Blind Date provided accurate dates and uncertainties that are consistent with our previous findings: all estimated dates lay within our uncertainty bounds, the accuracy generally depended on the resolution and quality of each image, and most estimated dates (for all sufficiently high-resolution images) were within a few years of the true dates. Our ground-truth dates are, unfortunately, very unreliable, as the EXIF data may simply reflect the date that an image was digitized or modified, rather than the date at which it was imaged. Although this means that we are not able to truly vet Blind Date's performance for these amateur images, this issue also highlights the utility of this system: the dates of origin of these images are effectively lost, but can be re-estimated.

5. DISCUSSION

We have shown that our Blind Date system can successfully attach time metadata to historical imaging data. The system runs in seconds on standard inexpensive consumer computer equipment; it does not require large investments of time or money to vet or create time metadata for large collections of astronomical imaging.

The performance of Blind Date will depend on the properties of the input image, and on the properties of the catalog
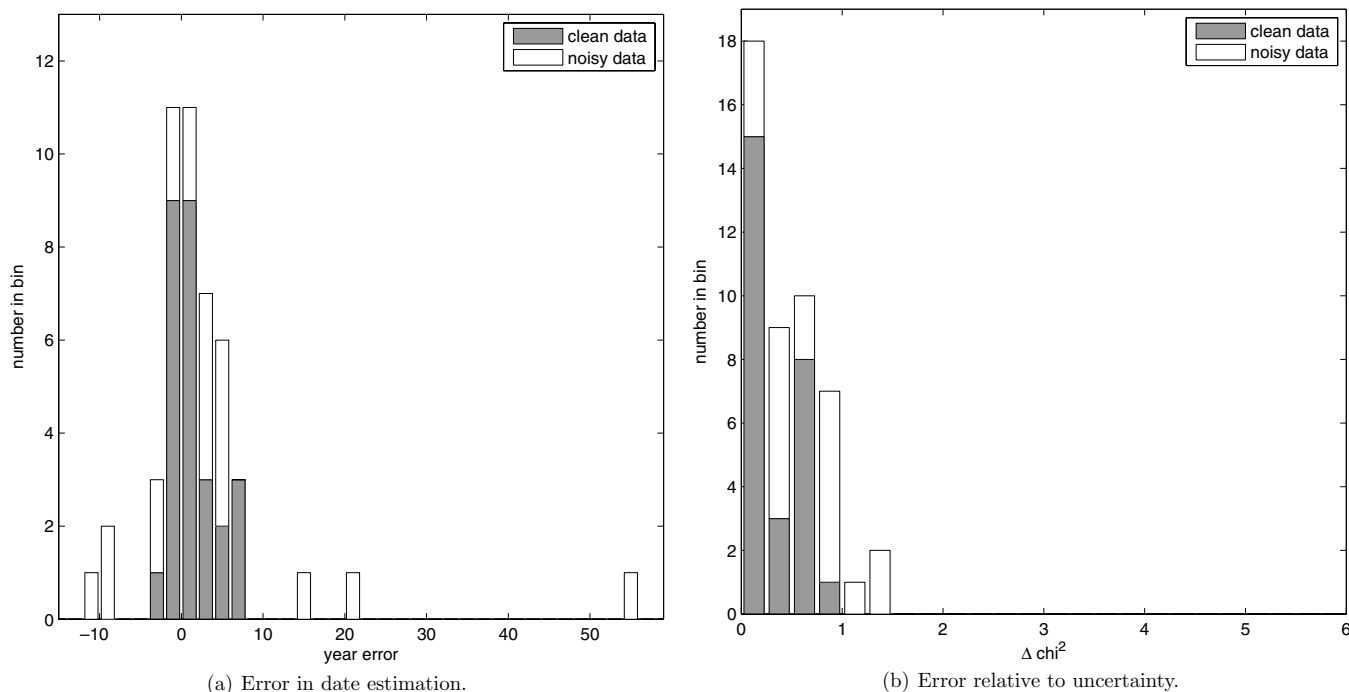
(a) Error in date estimation.   (b) Error relative to uncertainty.

Figure 7. Histograms of the errors in date estimation, for both data sets. (a) The difference between our estimated dates and the true dates is shown. (b) Those errors relative to the widths of the uncertainties are shown; effectively, the differences between the χ^2 scores of the true years and the χ^2 scores of the estimated years. A relative error of less than 1 indicates that the true year lies within the uncertainty. (a) Error in date estimation. (b) Error relative to uncertainty.

information known about the region of the sky that is being imaged. We can phrase this as two questions: "What is the information content in an image?" and "What is the information content in a catalog star?"

In an attempt to empirically assess the information content of an image, we ran a simple experiment in which we varied the resolution of an input image and the number of stars contained in the input image (by downsampling and cropping the image, respectively). The results are shown in Figure 8, where we see that the performance depends heavily on an image containing a large number of stars imaged at high resolution. We also explored the effects of different kinds of imaging defects. Our objective function is designed to be robust to false sources, and as such, Blind Date performs very well on images with multiple exposures. Our experiments suggest that saturation, large point-spread function (PSF) due to poor focus or trailing, and short exposure time (low sensitivity) most negatively affect Blind Date's performance. See Figure 5 for examples of our accuracy in the face of different kinds of imaging defects. Trailing and saturation can effectively be thought of as decreasing the resolution of our input image (by decreasing our ability to accurately determine the centroid of the stars), and shallow imaging is effectively equivalent to dropping dim stars out of the image; these are the two trends demonstrated in Figure 8. Once again, Blind Date's accuracy appears to depend on the number of well-imaged (high-resolution) stars in an image.

The information in a single catalog star (provided that it has been detected in the input image, in the limit that our procedure is equivalent to least-square fitting) is proportional to that star's contribution to the second derivative of χ^2 with respect to date. This contribution is approximately the square of the magnitude of the star's proper motion, divided by the square of the uncertainty, that is, the square of the S/N at

which the proper motion is detected (where the "noise" in this case is the combined uncertainty from the catalog and the image as in Equation (3)). We expect Blind Date's performance on an image to scale roughly with the sum of the squares of the detected catalog stars' proper motion S/Ns. Imagery unlike those analyzed here ought to obtain date calibration with uncertainty that goes down as the sum of the detected catalog stars' proper motion S/Ns goes up.

Increasing the polynomial order of the transformation on the image plane produces slightly more accurate date estimates, presumably because the input images do have distortions that are represented reasonably by these functions. We are reluctant to advocate unnecessarily large polynomial orders, as—theoretically—the more freedom we give to the transformation model, the more irregular our resulting χ^2 curves may become. This being said, we have not seen any evidence that reinforces such a concern. In principle, even more accurate results could be obtained without increasing the number of degrees of freedom in the fit by employing a physical camera model that represents known distortions in the particular camera used to take the data.

For each image in our data set, we performed an experiment to determine the range of initial dates for which our search algorithm is robust. We discovered that each image had a range of at least 300 years (and on average, 620 years), roughly centered around the true year from which the search could be initialized without the final result being affected. If we ignore our precaution of using gradient descent when the second derivative of the χ^2 curve is nonpositive, this range is significantly smaller (on average, about 55 years). This finding highlights the importance of the modifications we make to Newton's method in constructing our search algorithm, and also suggests that the coarse grid of queries with which we initialize our search is unnecessary—a single initial query in the correct century would have been more than sufficient.
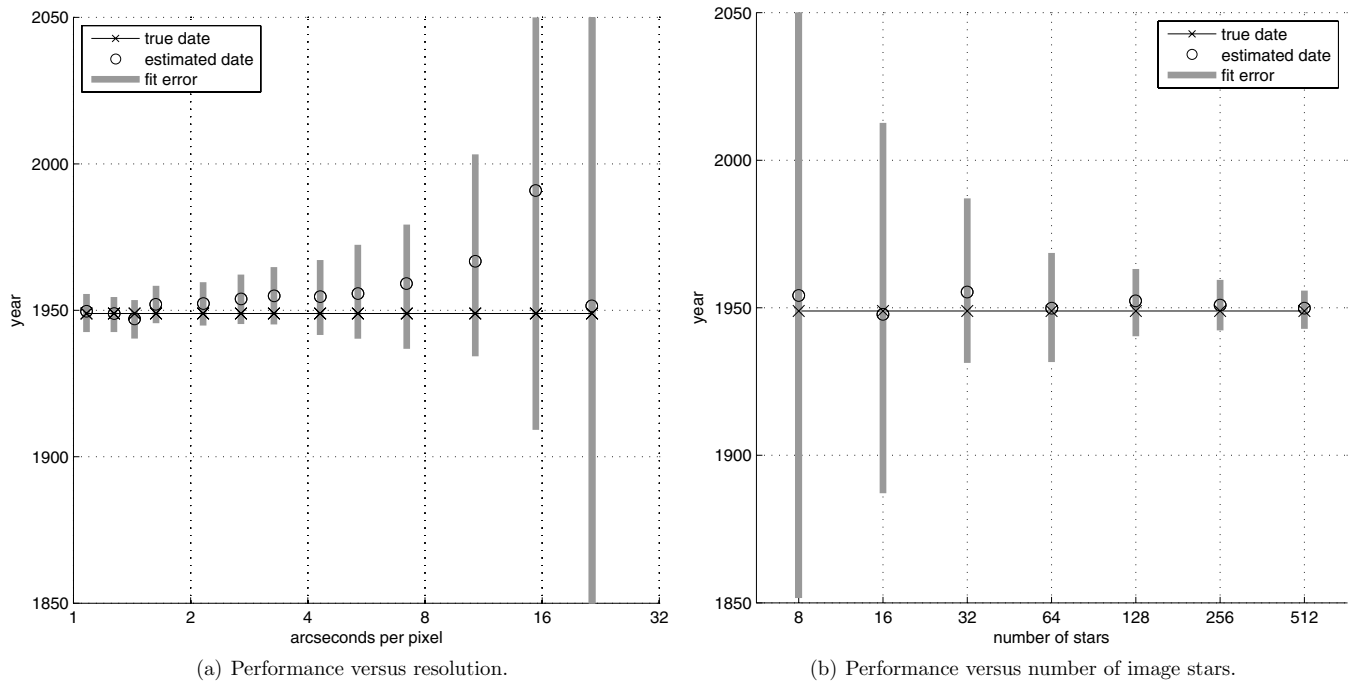
(a) Performance versus resolution.

(b) Performance versus number of image stars.

Figure 8. Plots showing performance relative to the resolution of the input image and to the number of stars the input image contains. (a) was produced by repeatedly downsampling the input image, while (b) was produced by repeatedly cropping out the borders of the input image. Note that in downsampling the image, some smaller stars are no longer detectable by our source-extraction algorithm, so the number of stars in the image decreases with the resolution of the image. (a) Performance vs. resolution. (b) Performance vs. number of image stars.

Blind Date largely ignores one very important source of information, namely the brightnesses of the image and catalog stars. These data are used in our bandpass estimation step, but are then largely ignored. In future versions of *Astrometry.net*, we plan to utilize these data in a number of ways. We eventually hope to simultaneously estimate all parameters of a given image, including the image's WCS, its date of origin, and its bandpass. Estimating all of these simultaneously should mean that brightness information is implicitly used in our estimation of the location and date, and should improve our results accordingly. This would also solve our current conundrum regarding this process, which is that image–catalog correspondences are required for bandpass estimation, while bandpass estimation is required for finding image–catalog correspondences.

Analysis of the brightness of image stars may be able to play a profound role in date estimation if we consider a subset with periodic variability. Given an image containing k stars with different periods, and given sufficient information concerning the periodic variations in their brightnesses, we should be able to constrain the date of origin of the image to one of a set of time intervals in which those k stars are at the correct points in their periods (to within photometric precision). Given the set of intervals constrained by the periodic variations, we can use the range of dates determined by Blind Date (i.e., from the proper motions of the stars) to select a potentially very narrow time interval in which the image must have originated. In principle, it may even be possible to determine the date of origin of an image solely through periodic brightness, though that would require very good measurements of the periods of the catalog stars and of the brightnesses of the image stars.

Blind Date's value, on the most superficial level, is to recover lost metadata (at low precision) for historical and amateur data that have been archived poorly or not at all. Now that large scanning projects are underway at photographic archives and the web is providing new opportunities for file sharing

among amateurs and professionals, there is a great need for systems that automatically vet and provide metadata for data of unknown provenance. In vetting data of good provenance, the Blind Date system makes historical data more reliable; in time-calibrating data of poorer provenance, the Blind Date system can recover "lost" data in the time domain. This recovery will improve measurements of time-dependent effects such as stellar variabilities, transients, stellar proper motions, and minor planet orbits, all of which can have uncertainties limited by the number of known historical images.

Regardless of whether or not an image already contains reliable date metadata, the techniques described in Blind Date may have implications for its calibration. The fact that the date can be estimated from input images demonstrates both that the images contain important information about stellar motions *and* that astrometric calibration will be hampered if it is performed with a catalog projected to an epoch far from the date that the catalog imagery was taken. A system that is time sensitive, such as Blind Date, will plausibly provide the best astrometric calibration possible for arbitrary imaging.

What may be Blind Date's most important consequence is an inversion of the system, in which we attempt to use imagery to re-estimate the proper motions of catalog stars. The most straightforward approach to this would be to "cheat" and use the ground-truth dates of all input imagery. One could repeatedly calibrate each image using the catalog that is wound to that image's date of origin, re-estimate the proper motions of the catalog stars, and rewind the catalog using those new proper motions. This could be thought of as performing expectation-maximization on the proper motions of the catalog. Of course, an ideal system would be robust to some (or all) input imagery not having ground-truth ages. We could estimate the date of origin of all unlabeled imagery, and use these estimates (and their uncertainties) in our expectation-maximization. Labeled and unlabeled data could be treated

equivalently, except that labeled data would have much less uncertainty associated with it. This system would then become a two-way street, in which we reposition not only images relative to the sky, but also the sky relative to the images, and dynamically develop a consensus between the two. The future *Astrometry.net* "catalog" would not have to be a static entity, but would instead be a consensus of all available imagery—using the USNO-B Catalog as a static "prior." Blind Date takes us one step closer to this grand long-term hope for *Astrometry.net* becoming an always-changing database of everything we know about the sky, by allowing *time* to become one more dimension of our data.

REFERENCES

Barron, J. T., Stumm, C., Hogg, D. W., Lang, D., & Roweis, S. 2008, AJ, 135, 414
Hampel, F. R., Ronchetti, E. M., Rousseeuw, P. J., & Stahel, W. A. 1986, Robust Statistics: The Approach Based on Influence Functions (New York: Wiley)
Monet, D. G., et al. 2003, AJ, 125, 984
York, D., et al. 2000, AJ, 120, 1579