A NEW METHOD FOR ROBUST HIGH-PRECISION TIME-SERIES PHOTOMETRY FROM WELL-SAMPLED IMAGES: APPLICATION TO ARCHIVAL MMT/MEGACAM OBSERVATIONS OF THE OPEN CLUSTER M37

S.-W. Chang1, Y.-I. Byun2, and J. D. Hartman3

1 Institute of Earth-Atmosphere-Astronomy, Yonsei University, Seoul 120–749, South Korea; seowony@galaxy.yonsei.ac.kr
2 Department of Astronomy and University Observatory, Yonsei University, Seoul 120–749, Korea; ybyun@yonsei.ac.kr
3 Department of Astrophysical Sciences, Princeton University, Princeton, NJ 08544, USA

ABSTRACT

We introduce new methods for robust high-precision photometry from well-sampled images of a non-crowded field with a strongly varying point-spread function. For this work, we used archival imaging data of the open cluster M37 taken by MMT 6.5 m telescope. We find that the archival light curves from the original image subtraction procedure exhibit many unusual outliers, and more than 20% of data get rejected by the simple filtering algorithm adopted by early analysis. In order to achieve better photometric precision and also to utilize all available data, the entire imaging database was re-analyzed with our time-series photometry technique (Multi-aperture Indexing Photometry) and a set of sophisticated calibration procedures. The merit of this approach is as follows: we find an optimal aperture for each star with a maximum signal-to-noise ratio and also treat peculiar situations where photometry returns misleading information with a more optimal photometric index. We also adopt photometric de-trending based on a hierarchical clustering method, which is a very useful tool in removing systematics from light curves. Our method removes systematic variations that are shared by light curves of nearby stars, while true variabilities are preserved. Consequently, our method utilizes nearly 100% of available data and reduces the rms scatter several times smaller than archival light curves for brighter stars. This new data set gives a rare opportunity to explore different types of variability of short (∼minutes) and long (∼1 month) time scales in open cluster stars.

Key words: methods: data analysis – open clusters and associations: individual (M37) – stars: variables: general – techniques: photometric

1. INTRODUCTION

We are poised on the threshold of unprecedented technical growth in wide-field time domain astronomy, where ground-based observations yield very precise measurements of stellar brightness from high-volume data streams. So far, wide-field time-series surveys have been spearheaded by relatively small telescopes since they are supported by large field of view (FOV) instruments operating with high duty cycle (see Becker et al. 2004 for a summary of optical variability surveys). Within the last decade, the advent of large mosaic CCDs has facilitated the coverage of large sky area even for large-aperture telescopes (e.g., MMT Megacam: McLeod et al. 2000; ESO Very Large Telescope Omegacam: Kuijken et al. 2002; Subaru Suprime-Cam: Miyazaki et al. 2002; CHFT Megacam: Boulade et al. 2003; iPTF: Kulkarni 2013). Although these facilities are generally devoted to imaging surveys, researchers are attempting to utilize them for short- and long-term variability surveys with short-cadence exposures (e.g., Hartman et al. 2008a; Pietrukowicz et al. 2009; Randall et al. 2011). Such wide-field imaging systems have enabled us to observe hundreds of thousands of target stars simultaneously and also to detect various variability phenomena. A remarkable thing about these surveys is that the fraction of variable sources increases as the photometric precision of the survey improves. For this reason, it is important to improve the accuracy in photometry.

Another key issue in wide-field time-series photometry is the removal of temporal systematics from a single image frame or several consecutive image frames. It has recently become known that systematic trends in time-series data can be different and localized within the image frame when the FOV is large. Such spatially localized patterns may be related to subtle point-spread function (PSF) differences and sky condition within the detector FOV (e.g., Kovács et al. 2005; Pepper et al. 2008; Bianco et al. 2009; Kim et al. 2009). As these patterns change in time, we can see how the temporal variations of systematic trends affect the brightness and shape of light curves directly. The time-scale of systematic variation is sometimes comparable to short-term variability, such as transits or eclipses, and in some cases even long-term variability. Thus, it is often difficult to identify and characterize true variabilities.

In this paper, we introduce a new photometry procedure, called multi-aperture indexing, which is suited to analyzing well-sampled wide-field images of non-crowded fields with a highly varying PSF, such as those produced by wide-field mosaic imagers on large telescopes. We apply this procedure to archival imaging data from the MMT/Megacam transit survey of the open cluster M37 (Hartman et al. 2008a), demonstrating a substantial improvement over the existing photometry.

Section 2 describes the MMT imaging database and identifies problems in the existing photometry which motivated the development of our new methods. Section 3 describes the multi-aperture photometry that utilize newly defined contamination index and carefully tuned calibration procedures, including the results of the basic tests to validate our approach. Section 4 gives an in-depth discussion about systematic trends in time-series data and suggests an efficient way for identifying, measuring, and removing spatio-temporal trends. Section 5 describes the effects of new calibration on period search, and we summarize our main results in the last section.

2. ARCHIVAL MMT TIME-SERIES DATA OF M37

Hartman et al. (2008a) have conducted a study to find Neptune-sized planets transiting solar-like stars in the rich open
cluster M37. The observing strategy was carefully designed for a transiting planet search by several considerations (e.g., the reliability of exposure time per frame, the effects of pixel-to-pixel sensitivity variations, and sensitivity of filter). Their work did not reveal any transiting planets, but it did provide a rare opportunity to explore photometric variability at relatively high temporal resolution with 30–90 s. Hartman et al. (2008b) discovered 1430 new variable stars, including very short-period eclipsing binaries (e.g., V37, V706, V1160) and δ Sct-type pulsating stars (e.g., V397, V744, V1412).

We used the same data set on the open cluster M37. A detailed discussion of the observations, original data reduction, and light curve production is described in Hartman et al. (2008a, 2008b). The data archive consists of approximately 5000 r'-filter images taken over 24 nights with the wide-field mosaic imager (Megacam) mounted at the f/5 Cassegrain focus of the 6.5 m MMT telescope. Note that Megacam is made up of 36 2048 × 4608 pixel CCD chips in a 9 × 4 pattern, covering a 24' × 24' FOV (McLeod et al. 2000). This instrument has an unbinned pixel scale of 0.08", but it was used in 2 × 2 binning mode for readout.

The observation logs are summarized in Table 1 of Hartman et al. (2008a). In brief, the r'-bandtime-series observations were undertaken between 2005 December 21 and 2006 January 21, with a median FWHM of 0.89 ± 0.39 arcsec. Exposure times are chosen to keep an r ∼ 15 mag star as close to the saturation limit, which is expressed as a function of seeing conditions. With an average seeing ∼0.089 on images, the quality of the images is good to achieve high-precision light curves (less than 1% rms value) down to 20. In addition to the imaging data set, this database includes light curve data sets for a total of 23,790 sources detected in a co-added reference image. Theses light curves are obtained by the image subtraction technique using a modified version of ISIS software (Alard & Lupton 1998; Alard 2000).

Figure 1. Example of MMT light curves for the brighter stars in archive. The dashed lines are weighted spline approximation with ±3σ control limits (dotted lines). These light curves contain outlier points that significantly increase the rms scatter of the raw light curves.

Figure 2. Tendency of optimal aperture selection from the multi-aperture photometry. As stars get fainter, the optimal aperture decreases in size. Sources outside of this sequence are suspected to be contaminated.
As shown in Figure 1, however, the raw light curves from the original image subtraction procedures exhibit many unusual outliers, and more than ∼15% of data get rejected by a simple filtering algorithm after cleaning procedures. In practice, brutal filtering that is often applied to remove outlying data points can result in the loss of vital data, with seriously negative impact to short-term variations such as flares and deep eclipses. We also find that the image subtraction technique often resulted in measurement failures from several frames due to poor seeing or tracking problem. After removing these bad frames, it leads to loss of additional ∼5% data points from most light curves. In order to overcome this problem, we have re-processed the entire image database with new photometric reduction procedures.

3. NEW PHOTOMETRIC REDUCTION

3.1. Preparation for Photometry

We followed the standard CCD reduction procedures of the bias correction, overscan trimming, dark correction, and flat-fielding as described in Hartman et al. (2008a). The individual CCD frames were calibrated in IRAF, using the mosaic data reduction package MEGARED.3 The first step is to correct the pixel-to-pixel zero-point differences that are usually described by the sum of a mean bias level and a bias structure. As the bias frames were not separately taken during the time of the observations, the mean bias level was subtracted from each image extension using an overscan correction and so we cannot remove any remaining bias structure from all overscan-subtracted data frames. According to description in Matt Ashby’s Megacam reduction guide,4 the bias structure can be very significant in some small regions such as the portions of the arrays close to the readout leads. The dark currents are normally insignificant for Megacam so that corrections are not needed even for long exposures. The next step is to correct pixel-to-pixel variations in the sensitivity of the CCD; we used the program domegacamflat2.5 This program determines the scaling factors to correct the gain difference between the two amplifiers of each chip by finding the mode in the quotient of pixels to the left and right of the amplifier boundary, and then flattens each of the frames with a master flat-field frame. It is worth mentioning that the sky conditions were rarely photometric during the observing run, with persistent light cirrus for most of the nights. Therefore it was only possible to obtain

3 The 64 bit version of Megacam reduction package is available from https://cfa.harvard.edu/~bmcleod/Megared/.

4 The detailed reduction procedures of MMT/Megacam are described by Matthew L. N. Ashby at https://cfa.harvard.edu/~mashby/megacam/megacam_frames.html.

5 This C routine is written by J. D. Hartman to reduce I/O overhead. It reads in a list of Megacam mosaic images and a Megacam mosaic flat-field image, and then writes out the flattened image over the existing image.
twilight sky flats on a handful of nights (dome flats were not possible).

We removed bad pixels using the Megacam bad pixel masks distributed with the MEGARED package. The values of bad pixels are replaced with interpolated value of the surrounding pixels using the IRAF task fixpix. The numerous single pixel events (cosmic-rays) were identified and removed using the LACosmic package (van Dokkum 2001).

The Megacam data already have a rough World Coordinates System (WCS) solution that is based on a single value of the telescope pointing. To update these with a more precise solution, we applied astrometric correction to each CCD in the mosaic using the WCSTools imwcs program (Mink 2002). The new solution is derived by minimizing the differences between the R.A. and decl. positions of sources in a single CCD chip and their positions listed in the 2MASS Point Source Catalog (Skrutskie et al. 2006). The resulting astrometric accuracy is typically better than 0\".1 rms in both R.A. and decl.

### 3.1.1. Building a Master Source Catalog

Typically, a point or extended source detection algorithm is applied to each frame independently and it always requires criteria for what should be regarded as a true detection. In obtaining the pixel coordinates for all objects in the M37 fields, this procedure often misses some objects when the detection threshold approaches the noise level. Also it needs a substantial effort to match the objects that are detected in only some of the frames. Our approach is as follows: a complete list of all objects is obtained from a co-added reference frame, and then the photometry is performed for each frame using the fixed positions of the sources detected on the reference. Since the relative centroid positions of all objects are the same for all frames in the time series, we can easily place an aperture on each target and measure the flux even for the stars at the faint magnitude end.

We constructed the reference frame for each chip from the best seeing frames using the SWarp software. Benefiting from a highly accurate astrometric calibration of input frames, we were able to improve the quality of co-added images. In the SWarp implementation, the pixels of each frame were resampled using the Lanczos3 convolution kernel, then combined into the reference frame by taking a median or average. After this was done, sources were detected and extracted on the reference frame using the SExtractor software (Bertin & Arnouts 1996). When configured with a lower detection threshold, SExtractor extracts the number of spurious detections (e.g., diffraction spikes around bright stars, or outer features of bright galaxies). These false detections were removed by careful visual inspection for each chip. The final catalog contains a total of 30,294 objects including both point and extended sources.

---

6 SWarp is a program that resamples and co-adds FITS images, distributed by Astromatic.net at [http://astromatic.net/software/swarp](http://astromatic.net/software/swarp).
3.1.2. Refining the Centroid of Each Object

Prior to the photometry, the initial centroid coordinates of the target objects for each frame were computed by using the WCSTools sky2xy routine (Mink 2002). The stored world coordinate system for each frame is used to convert the (R.A., decl.) coordinates from the master source catalog to the (x, y) pixel locations. However, actual positions of objects for each frame can be slightly moved from its original locations depending on the focus condition of instrument, seeing condition, and the signal-to-noise ratio (S/N) in the individual observations. These types of positioning errors (i.e., centroid noise) will lead to the internal error for a photometric measurement that results from placement of the measuring aperture on the object being measured. The situation gets worse for faint stars because the centroid position of them is itself subject to some uncertainty. The fractional error in the measured flux as a result of mis-centering is given by:

$$\frac{\delta F}{F_0} \approx \frac{1}{\sqrt{2\pi}} \frac{\Delta}{\sigma} e^{-\frac{2\Delta^2}{\sigma^2}}$$

where $\Delta$ is the positioning error, $R$ is the radius of aperture, and $\sigma$ is the profile width for a source with a Gaussian PSF and total flux $F_0$ (see Appendix A in Irwin et al. 2007 for details). This expression shows that if we set the aperture radius equal to the

![Figure 5. Histogram of differences in the magnitude offsets between the left- and right-side of each chip ($\Delta_l - \Delta_r$) for the whole data set. Each histogram is normalized by the total number of data frames $N$ and is described by Gaussian distribution with different means and variances (dashed lines).](image-url)
FWHM \( (R = 2.35\sigma) \), even small differences in placement of the aperture \( (e.g., \Delta = 0.1\sigma) \) may increase the uncertainty in the flux measurements \( (\approx 1\altmag) \). Thus, accurate centroid determination is important to achieve the high-precision photometry.

Following the windowed centroid procedure in the SExtractor, a refined centroid of each object is calculated iteratively. On average, the rms uncertainty in the coordinate transformation using the WCS information was \( 0.044 \pm 0.01 \) pixels for the bright reference stars. After the coordinate transformation from sky to \( xy \), however, the centroid coordinates \( (x_{\text{final}}, y_{\text{final}}) \) are slightly misaligned from their actual ones \( (x_{\text{center}}, y_{\text{center}}) \). The refined centroid values \( (x_{\text{final}}, y_{\text{final}}) \) are used only if the maximum displacement is at least less than 1.5 pixels. This condition prevents arbitrary shifting of a source centroid, especially for faint stars.

### 3.1.3. Estimation of Background Level

We estimated a local sky background by measuring the mode of the histogram of pixel values within a local annulus around each object, which is suitable choice for our uncrowded field \( (\text{less than } \sim 1000 \text{ stars per chip}) \). This process is a combination of \( \kappa-\sigma \) clipping and mode estimation. The background histogram is clipped iteratively at \( \pm 3\sigma \) around its median, and then the mode value is taken as:

\[
\text{Mode} = 3 \times \text{Median} - 2 \times \text{Mean}.
\]  

It represents the most probable sky value of a randomly chosen pixel in the sample of sky pixels \( (\text{Stetson 1987}) \).

### 3.2. Multi-aperture Indexing Photometry

Modern data reduction techniques aim to reach photon noise limit and minimize systematic effects. For example, differential photometry technique can be achieved better than 1% precision for brighter stars \( (e.g., \text{Everett & Howell 2001}; \text{Hartman et al. 2005}) \), and the deconvolution-based photometry algorithm leads to the minimization of systematic effects in very crowded fields \( (e.g., \text{Magain et al. 2007}; \text{Gillon et al. 2007}) \). However, conventional data reduction methods often fail to handle various artifacts in wide-field survey data. We present below a new photometric reduction method for precise time-series photometry of non-crowded fields, without the need to involve complicated and CPU intensive process \( (e.g., \text{PSF fitting or difference image analysis}) \).

#### 3.2.1. Photometry with Multiple Apertures

Our photometry is similar to standard aperture photometry, except in that we compute the flux in a sequence of several apertures and then determine the optimum aperture individually to each object at each epoch. This multi-aperture photometry is an efficient way to determine the optimum aperture size that gives the maximum S/N for a flux measurement. The maximum S/N is not necessarily at the same aperture for all objects, and it can be obtained from a relatively small aperture \( (\text{Howell 1989}) \). This photometric aperture is to achieve the optimal balance between flux loss and noises based on a relationship derived from the CCD equation \( (\text{see Merline & Howell 1995}) \). Figure 2 shows how the optimum apertures vary with the stellar magnitude. There is an obvious trend of decreasing aperture sizes with increasing magnitudes down to the faint magnitude limit in the example frame.

Once we measure the flux of each object with the optimum aperture, we need to apply the aperture correction for small apertures. The aperture correction terms are estimated from the growth curve analysis of selected isolated bright stars \( (i.e., \text{reference stars}) \). The average curve-of-growth for each frame is calculated by measuring the difference in magnitude between different pairs of apertures \( (\text{up to 10 pixels aperture radius}) \) and then an automatic correction is applied to all objects for each photometric aperture. The use of a common aperture correction for each CCD assumes that there is no variation in the correction across the CCD. This flux correction method gives nearly the same brightness within the measurement uncertainties for all apertures. Any PSF variation across the CCD causes systematic errors, however, and we deal with this in Sections 3.4 and 4.

We performed the multi-aperture photometry based on the concentric aperture photometry algorithm in DAOPHOT package \( (\text{Stetson 1987}) \), using several circular apertures \( (\text{up to 10 pixels aperture radius}) \) with a fixed sky annulus from 35
to 45 pixels. The initial results of multi-aperture photometry are stored in ascii-format photometry tables, including the date of the observations (MJD), the pixel (x, y) coordinates, the aperture-corrected magnitudes with errors for each aperture, the sky values and its errors. Figure 3 shows the details of the multi-aperture photometry for one star at different epochs. In the former two epochs, the photometric apertures can be properly selected by the S/N cuts, while in the latter two epochs, S/N increases for larger apertures. This unusual behavior is due to contamination by a moving object. We automatically identifies similar unusual cases by the method of aperture indexing.

3.2.2. Determination of the Best Aperture with Indexing Method

Our multi-aperture indexing method is similar to the basic concept of the discrete curve-of-growth method (Stetson 1990). Each object is indexed based on the difference in aperture-corrected magnitude between pairs of apertures \( \Delta m_1 = \hat{m}_{i,j,ref} - \hat{m}_{i,j,k} \) with mean trend for stars of similar brightness (see solid and dotted lines in bottom panels of Figure 4, respectively). The aperture with a 10 pixel radius is used as the fixed reference aperture. The mean trend is determined by computing the rms curve of the aperture correction values for all measured apertures, and used to evaluate whether magnitude at a given aperture significantly differs from the mean trend. Since the rms value depends on the chosen magnitude interval, all stars are divided into groups according to their brightness in the individual frames. We determine the rms curve for each magnitude group using an iterative \( \sigma \)-clipping until convergence is reached. Objects lying within \( \pm 3 \sigma \) of the model curve are indexed as contamination-free, and those above \( \pm 3 \sigma \) as a contaminated source. Figure 4 shows that multi-aperture indexing guides us to throw out some photometric measurements if they are discrepant from the mean trend. This approach also gives us a chance to recover a measurement that would be otherwise thrown out. The problematic aperture can be simply replaced by one of the smaller apertures if it is indexed as contamination-free. This help us make a full use of the information offered by the data.

3.3. Improved Photometric Calibration

We present a new photometric calibration to convert the instrumental magnitudes onto the standard system, including a relative flux correction of the left and right half-region of each CCD chip. As mentioned in the Section 3.1, MMT/Megacam shows the temporal variations in the gain between two amplifiers on each CCD, as well as between CCDs that are part of the same mosaic. It may have been caused by unstable
bias voltage of the CCD output drain which has a profound impact on the gain of the output amplifier. The level of readout noise is also unstable between two amplifiers. To correct for this effect, the photometric calibration needs to be performed individually for each amplifier region.

We use a sufficient number of (pre-selected) bright isolated stars as standard stars and compute the relative flux correction terms. These terms were derived for each frame using the mean magnitude offset \( \Delta m_{ij} \) of standard stars \( (N_{lr},) \) with respect to corresponding magnitudes \( (m_{ij},) \) in the master frame chosen as an internal photometric reference

\[
\Delta m_{ij} = \frac{1}{N_{lr}} \sum m_{ij} - \bar{m}_{ij} + ZP_{lr},
\]

where \( \bar{m}_{ij} \) is the aperture-corrected instrumental magnitudes in other frames and \( ZP_{lr} \) is the photometric zero-points for the left- and right-side of each chip, respectively. To calculate the zero-points, we solve a linear calibration relation of the form:

\[
r - m_{ij} = ZP_{lr} - 0.07X + 0.107(r - i),
\]

where \( r \) and \( i \) are standard magnitudes from the photometric catalog of M37 (Hartman et al. 2008a) and \( X \) is an airmass term. The fit is performed iteratively using a sigma-clipping method.

Figure 5 shows the difference in the magnitude offsets between the left- and right-side of each chip \( (\Delta_l - \Delta_r) \) for the whole data set, which is within \( \pm 0.02 \) magnitude level for all 36 CCD chips. The histograms are normalized by the total number of data frames \( N = 4730 \) and are described by a Gaussian function with slightly different mean values and shapes (dashed line). We clearly see a significant variation in difference between a pair of magnitude offsets for all CCDs.

### 3.4. Field Distortion Correction

The photometric calibration for wide-field imaging systems is also affected by position-dependent systematic errors due to a PSF variation across the FOV (e.g., Ivezić et al. 2007; Hodgkin et al. 2009). We derive PSF variations across the FOV with the SExtractor package. The change of the PSF shapes in the image plane is represented by spatial distribution of PSF FWHM values for several bright stars, with parameters of \( \text{CLASS~STAR} > 0.9, \text{MAGERR~AUTO} < 0.01 \) mag, and \( \text{FLAGS} = 0 \) (i.e., isolated point sources with no contamination). Note that the PSF FWHM values are defined as the diameter of the disk that contains half of the object flux based
on a circular Gaussian kernel. Figure 6 presents the variation of the PSF FWHM as a function of distance from the image center for various seeing conditions. For each quadrant, the dashed lines represent the weighted spline approximation of the median value of each distance bin (1 arcmin). The result shows that the PSF FWHM varies significantly as a function of position on the single-epoch image frames and variations are at the level of ∼10% to 20% (0"/1 – 0"/2) across the FOV. As the field distortion is not negligible from the center of field to its edges, such variations limit the accuracy of stellar photometry.

To address this issue, we perform a 2D polynomial fitting technique. For each frame, the correction terms are described by a linear or quadratic polynomial depending on the position \((x, y)\) only.

\[
\Delta m_{c,k}(x, y, m_{\text{err}})_{\text{lin}} = c_{0k} + c_{1k}x + c_{2k}y,
\]

\[
\Delta m_{c,k}(x, y, m_{\text{err}})_{\text{quad}} = c_{0k} + c_{1k}x + c_{2k}y + c_{3k}x^2 + c_{4k}y^2 + c_{5k}xy,
\]

where \(x, y\) are the pixel coordinates of \(N\) bright isolated stars, \(m_{\text{err}}\) is the statistical weight in the fitting procedure, \(c_k\) are the sets of polynomial coefficients for each aperture size, and \(\Delta m_{c,k}(x, y, m_{\text{err}})\) are the difference in magnitude between the reference aperture and \(k\) aperture, \(\Delta m_{c,k}(x, y) = m_{c,20}(x, y) - m_{c,k}(x, y)\), at the position \((x, y)\) for each chip \(c\). We derived the optimal parameter values from a nonlinear least-squares fit using the Levenberg–Marquardt algorithm and automatic differentiation,\(^7\) and choose between two models that best fit the data.

Figure 7 shows the field-dependent magnitude offsets and the distortion correction by 2D polynomial fitting method for one example mosaic CCD. For the outer and the central region of the mosaic, we compare the magnitude offsets between the reference aperture and the relatively smaller apertures as a function of \((x, y)\) coordinates. Here \(x\)-axis is in the declination direction and \(y\)-axis is opposite to the right ascension direction. We find that the magnitude difference depends on position \((x, y)\) and is most discrepant in the outer part of the FOV. This effect is usually more significant in the \(y\)-direction than in the \(x\)-direction, especially for the case of aperture photometry performed with small apertures. The correction for field-dependent PSF variation reduces the initial ∼10% variation (gray lines) to less than ∼1% (black lines).

\(^7\) We used a LeastSquareFit module provided in the scientific python package [http://scipy.org/].
3.5. Photometric Performance Diagnostics

3.5.1. Comparison with Archival Data

We compare the photometric performance of the recalibrated light curves with the non-de-trended archival light curves by means of the two representative measures: (i) the rms photometric precision $\sigma_{ph}$, and (ii) the data recovery rate $N_{\text{recovery}}$. The former is defined as the standard deviation of light curves around the mean value as a function of $r$ magnitude:

$$\sigma_{ph}(N_j, \bar{m}_j) = \sqrt{\frac{\sum_{p=1}^{N_j} (m_p - \bar{m}_j)^2}{N_j - 1}}, \quad (5)$$

where $N_j$ is the number of data points in each light curve, $m_p$ is the observed magnitude, and $\bar{m}_j$ is the mean magnitude of the object $j$, and the latter refers to the number of analyzed data frames normalized by the total number of observed data frames $N$ for each object. In typical cases, the data recovery rate should be near unity in the bright magnitude regime and decreases with magnitude for fainter objects. For comparison, we decided to select light curve samples which show either no significant variability or seeing-correlated variations induced by image blending. We remove all known variable stars from the sample list based on a new catalog of variable stars in M37 field (Chang et al. 2015, hereafter Paper II). To remove the light curves of blended objects, we use an empirical statistical technique to quantify the level of blending by looking for seeing-correlated shifts of the object from its median magnitude (Irwin et al. 2007).

$$b = \frac{\chi^2 - \chi^2_{\text{poly}}}{\chi^2}, \quad (6)$$

where

$$\chi^2 = \sum_p \frac{(\sum m_p - \bar{m}_j)^2}{\sigma_p^2} \quad (6)$$

for light curve points $m_p$ with uncertainties $\sigma_p$, and $\chi^2_{\text{poly}}$ is the same statistic measured with respect to a fourth order polynomial in FWHM fitted to the data. We adopt the value $b < 0.4$ for the selecting light curves with no blending. The last selection criterion is that the light curves must exit both in the archive and our database.

In the bottom panels of Figures 8 and 9, we plot the rms photometric precision of light curves for the two Megacam CCD chips in the outer (CCD 1) and central (CCD 21) part of the FOV, respectively. The black points show the rms values of the re-calibrated light curves, while the gray points are for the raw and filtered light curves in archive. The first impression from this comparison is that the typical rms scatter is

---

$^8$ Note that the archival data have not been trend-filtered, but Hartman et al. (2009) used the de-trended archival light curves as part of the transit detection process in their paper IV of the prior analysis.
overestimated from the raw light curves because of many outliers in the photometric data (bottom left panels). For the better results, these light curves were filtered out in two steps: (i) clipping 5σ outliers from each light curve and (ii) removing every data points that are outliers in a large number of light curves (Hartman et al. 2008b). In the second step, the outlier candidates are estimated by choosing a cutoff value for each CCD chip. The cutoff value is defined as the fraction of light curves for which a given image is a 3σ outlier. This filtering was applied to remove bad measurements due to image artifacts or poor conditions, which were previously thought to be unrecoverable, but resulting data loss is up to 20% of the total number of data points (bottom right panels). As shown in the top panels of the two figures, the data recovery rate for the recalibrated light curves is close to 100% over a wide range of magnitude and it appears to be more complete compared with the raw (top left panels) and filtered (top right panels) light curves. At bright magnitudes ($r < 16$) the data recovery rate does not reach 100% because the exposure time was chosen to be saturated at a magnitude of $r \sim 15$. This comparison proves that our approach is a powerful strategy for improving overall photometric accuracy without the need to throw out many outlier data points.

Finally, we compare the light curves themselves for selected variable stars between the archive and our own. This comparison serves to illustrate how the photometric precision and data recovery rate of the time-series data affect the ability to address a variety of variability characteristics. Figure 10 shows a direct comparison with the filtered light curves (top panels) and our re-calibrated light curves (bottom panels) for four variable stars. It is shown that our method recovers more data points (black) from the same data set of images.

3.5.2. Comparison with PSF-fitting Photometry

In order to further investigate possible systematics in our approach, we conducted PSF-fitting photometry with DAOPHOT II and ALLSTAR (Stetson 1987). For each mosaic frame, we select bright, isolated, and unsaturated stars to make the PSF model varying quadratically with ($x, y$) coordinates. After PSF modeling, we run ALLSTAR to perform iterative PSF photometry of all detected sources in the frame with initial

---

**Figure 11.** Example of residual magnitudes and sky values between our multi-aperture photometry and the PSF-fitting photometry as a function of $x$ and $y$ coordinates, respectively, for selected CCD chips. The filled squares are relatively bright stars with $r \leq 21$ mag, while the open squares are stars with $r > 21$ mag.
centroids set to the same values used for our own photometry. We then calculated aperture corrections using the package DAOGROW (Stetson 1990) after subtraction of all but PSF stars, which creates aperture growth curves for each frame and then integrates them out to infinity to obtain a total magnitude for each PSF star. The final step is to convert the instrumental magnitudes into the standard photometric system. For each frame, the initial zero-point correction is applied by correcting the magnitude offset with respect to the master frame. This places photometry for all frames on a common instrumental system. Following the same procedure in Section 3.3, the photometric calibration is performed individually for each amplifier region.

Figure 11 shows the residual magnitudes and sky values between our multi-aperture photometry and the PSF-fitting photometry as a function of position in the selected CCD chips. There are no position-dependent trends in the magnitude residuals. For the brighter stars with $r \leq 21$ mag, the rms magnitude difference between the two methods is very small ($D_{\text{chip1}} = -0.001 \pm 0.009$ and $D_{\text{chip21}} = 0.001 \pm 0.012$, respectively), while for the relatively faint stars the rms difference is somewhat larger ($D_{\text{chip1}} = -0.002 \pm 0.055$ and $D_{\text{chip21}} = 0.006 \pm 0.050$, respectively). The results of this example indicate that we can reliably correct for the PSF variations by our calibration procedures. Meanwhile, our sky values are slightly higher than the ALLSTAR sky values, but not to the degree that can seriously affect photometric measurements.

Figure 12 shows a comparison of the photometric precision (rms) for our multi-aperture photometry and for the PSF-fitting photometry as a function of the $r$ magnitude in the central region of the open cluster M37 (left panel). The arrows indicate the three different magnitude levels from bright to faint in our sample shown in the right panels. Note that the variable object IDs are taken from the new variable catalog of the M37 (see Paper II).

Figure 13. Example of shared systematic trend in the light curves of isolated bright stars. The numbers on the upper left- and right-side of each panel show the star’s identification number and the mean value of magnitude, respectively. The numbers on the x-axis indicate the corresponding timestamps in each frame. The deviations from the mean value, $\Delta r$, are less than $\pm 0.01$ mag level (with rms values of 0.0021, 0.0025, and 0.0033 mag from the top panel down). These stars show a similar pattern of light variations over the observation span.

The Astronomical Journal, 149:135 (19pp), 2015 April Chang, Byun, & Hartman
shown in the right panels of Figure 12, it is clear that our photometry tend to have smaller measurement errors with respect to the PSF photometry for the bright stars.

4. TEMPORAL TRENDS IN THE RE-CALIBRATED LIGHT CURVES

From a visual inspection of the re-calibrated light curves in the same CCD chip, we found that some light curves tend to have the same pattern of variations over the observation span (Figure 13). This kind of systematic variation (i.e., trend) is often noticed in other studies. For example, the importance of minimizing known (or unknown) systematics have been recognized by several exo-planet surveys because planet detection performance can be easily damaged by them (e.g., Kovács et al. 2005; Tamuz et al. 2005; Pont et al. 2006). Also space-based time-series data (e.g., CoRoT and Kepler) are no exception to this behavior although it is completely free from systematics caused by the turbulent atmosphere. Most of the raw light curves are affected by a secular (or a sudden) variation of flux without any obvious physical reason (Mazeh et al. 2009; Mislis et al. 2010; Jenkins et al. 2010).

In order to check the properties of temporal systematics, we examined the correlation coefficients as measure of similarity between two light curves $i$ and $j$ obtained from a single CCD chip (CCD 1).

$$C_{ij} = \frac{1}{N - 1} \frac{\sum_{t=1}^{N} (L_i(t) - \bar{L}_i)(L_j(t) - \bar{L}_j)}{\sigma_i \sigma_j},$$

where $L(t)$ is the flux of each star at time $t$, $N$ is the total number of measurements, $\bar{L}$ is the mean flux of each star, and $\sigma$ is the standard deviation of $L(t)$. This is a point-by-point comparison and is done for every pair of light curves in the data set. The resultant similarity matrix can be used to identify correlated pairs of light curves and determine which light curve is least like all other light curves (e.g., Protopapas et al. 2006). After that, we selected stars showing most systematics based on a hierarchical clustering method with the correlation coefficients (see Kim et al. 2009, for more details). Figure 14 represents spatial distribution of the most prominent trend groups on the CCD plane (top panel) and its strongly correlated features determined by the weighted sum of normalized light curves (bottom panel). There are two interesting features in this figure: the first one is that each trend covers only a certain part of the sky area and the second one is that some portions of neighboring trends show different variation patterns even at the same moment in time (shaded gray region in the figure). In

![Figure 14. Top panel: spatial distribution of the most prominent trend groups on the CCD plane which covers the small areas of the sky ($6.14' \times 2.73'$). The colored dots represent the most systematic stars classified as group 1 (red), group 2 (blue), and group 3 (green), respectively. One interesting property of trends is that they are localized within a CCD frame where stars are isolated from each other. Bottom panel: systematic features extracted from the selected light curves. These are calculated by the weighted sum of normalized light curves from each group.](image)
particular, we found an anti-correlated variation for the trends between the group 1 ($g_1$) and the group 2 ($g_2$), so we might expect to find possible noise sources that are responsible for these discrepancies. Why the trends are different and localized within a single CCD frame is a subject of further study, but it is probably related to subtle changes in point-spread function and sky condition within the detector FOV.

For these two groups, we consider a possible causal relationship between the systematic trends and average object/image properties. Figure 15 shows the differences in trend, differential magnitudes, sky level, and PSF FWHM between the two groups, respectively. In the top panel, we plot the magnitude difference in trends, which shows variations in the range of ±0.02 mag. We suspect that this may be due to the different concentration of star light between these two groups. It can be checked by using the magnitude difference $\Delta m = m_{20} - m_{ap}$, where $m_{20}$ and $m_{ap}$ are the reference aperture and the relatively small aperture, respectively. In fact, we already know that there is a magnitude variation in $\Delta m$ depend on the aperture size due to the field-dependent PSF variation (see Section 3.4). For example, Figure 16 shows the response of multi-aperture photometry for the two groups of stars at one epoch ($MJD = 53726.14817$) before and after applying the distortion corrections. Although the magnitude variation between the group 1 and the group 2 seems to have different behavior as a function of aperture size (gray points), it is negligible after the removal of the PSF variation (black points).

![Figure 15](image1.png)

**Figure 15.** Example subset of data for comparing the trends with several parameters from the two groups ($g_1$ and $g_2$) of stars. The numbers on the x-axis indicate the corresponding timestamps between the frame 1 and the frame 400. From the top to bottom, the panels show the average difference of trend, flux concentrations in which we applied a -0.02 mag shift to the $m_{20} - m_{12}$ value, sky level, and PSF FWHM value between the two groups.

![Figure 16](image2.png)

**Figure 16.** Systematic variations in the magnitude offsets, $\Delta m$, for the original multi-aperture magnitude measurements as a function of aperture size, which is marked with a gray dots in all panels. A positive $\Delta m$ indicates that the photometric measurements with the corresponding aperture size, $m_{ap}$, are brighter than those of reference aperture $m_{20}$, while a negative $\Delta m$ indicates vice versa. The dashed lines are the rms model profiles introduced in Section 3.2.2. There is a noticeable distinction between the group 1 (top panels) and group 2 (bottom panels) when looked at different concentration levels. But the trends are not fully explained by the difference patterns in $\Delta m$ because those correlated variations become small after the correction for the PSF variation (black dots).
We also check that the possible contribution of sky level ($\Delta$sky) and PSF FWHM differences ($\Delta$FWHM) to the systematic trends on the re-calibrated light curves. As mentioned by Bramich & Freudling (2012), sky over-subtraction may lead to the systematic trends as a function of the PSF FWHM, the amplitude of which increases for fainter stars. The third and forth panels of Figure 15 show the variation of the mean $\Delta$sky and $\Delta$FWHM, respectively. In our case, however, the form and amplitude of trends seem independent of sky level and PSF FWHM.

Some other possible sources that may contribute to the observed systematic trends include: higher order variations in the PSF shape beyond just the FWHM; cross-talk from other amplifiers; or ghosts from bright stars undergoing multiple reflections within the optics; non-uniform variations in the gain; and unmodeled temporal atmospheric variations that are dominated by Rayleigh scattering, molecular absorption by ozone and water vapor, and aerosol scattering (Padmanabhan et al. 2008). While we find a clear presence of trends that should be removed, we are not able to identify their exact cause.

4.1. Removal of Temporal Systematics by Photometric De-Trending (PDT)

In order to reduce systematic effects in photometric time-series data, several methods were introduced (e.g., TFA: Kovács et al. 2005; Sys-Rem: Tamuz et al. 2005; PDT: Kim et al. 2009; CDA: Mislis et al. 2010; PDC: Twicken et al. 2010). All of these algorithms share a common advantage that they work without any prior knowledge of the systematic effects. We use the PDT algorithm, which has been designed to detect and remove spatially localized patterns. By default, this algorithm works with a set of light curves that contain the same number of data points distributed in the same series of epoch. In many cases, however, missing data occur when no photometric measurements are available for some stars in a given observed frame. These missing data can be simply replaced by means, medians, or the values from the interpolation of adjacent data points in each light curve (e.g., Kovács et al. 2005; Kim et al. 2010). Although using the replaced value is the easiest way to reconstruct the light curve to be analyzed, it is not appropriate if the time separation between two subsequent observations is too large. Instead we use more
straightforward approach by applying the PDT algorithm in two separate steps: (i) we construct the master trends from the subset of bright stars, and (ii) de-trend light curves of all stars with most similar master trend and matching time line.

We briefly describe the main procedure of our de-trending process following the algorithm derived by Kim et al. (2009).

We first select the template light curves from bright stars that show the highest correlation in the light-curve features. The total length of template light curves should be long enough to cover the whole time span of observations. In this step, we take a sequence of data points, $L_i(t_{\text{ref}})$ as the reference time line. Using the correlation matrix calculated from Equation (8), we extract all subset of light curves that show spatio-temporally correlated features (i.e., clusters). Each cluster is determined by hierarchical tree clustering algorithm based on the degree of similarity. Next, we obtain master trends $T_c(t_{\text{ref}})$ for each cluster by weighted average of the normalized differential light curves, $f_i(t_{\text{ref}})$:

$$
T_c(t_{\text{ref}}) = \frac{\sum_{i=1}^{N_c} w_i f_i(t_{\text{ref}})}{\sum_{i=1}^{N_c} w_i},
$$

$$
f_i(t_{\text{ref}}) = \frac{L_i(t_{\text{ref}}) - \bar{L}_i}{\sigma_i},
$$

$$
w_i = \frac{1}{\sigma_i^2},
$$

where $N_c$ is the total number of light curves in each cluster $c$, $\bar{L}_i$ is mean value of $i$th light curve, and $\sigma_i$ is the standard deviation of $f_i(t_{\text{ref}})$. After determining the master trends, we de-trend the light curves of all stars with matching master trend and time line. We adjust the temporal sequence of measurements for the master trends $T_c(t_{\text{ref}})$ by that of individual light curves to be de-trended $L_i(t_i)$. Because each light curve is assumed as a linear combination of master trends and noise, we can determine the optimal solution by minimizing the residual between the master trends and the light curve:

$$
\hat{f}_i(t_i) = \sum_{c=1}^{m} \beta_{ic} T_c(t_i) + \epsilon_i(t_i),
$$

$$
\hat{f}_i(t_i) = \frac{L_i(t_i) - \bar{L}_i}{\sigma_i},
$$

where $m$ is the total number of master trends and $\beta_{ic}$ are free parameters to be determined by means of minimization of noise term $\sum_{c} \epsilon_i(t_i)^2$.

Figure 17 shows examples of our light curves before and after removing the systematic trends. The algorithm we used for de-trending removes only the systematic variations that are shared by light curves of stars in the adjacent sky regions (left panels), while all kinds of true variabilities are preserved (right panels).
5. IMPACT OF THE NEW CALIBRATION ON PERIOD SEARCH

The usefulness of our photometry is tested for a set of variable stars. We immediately find abundant cases of improvements in the following three aspects: (i) refinement of the derived period, (ii) detection of a new significant peak in the periodogram, and (iii) separation of non-variable candidates where systematics in the light curves were mistaken for true variability. For each case, we compare light curves and power spectra for archival data and our data.

For the first example, we show that a new photometric measurement and calibration allowed us to derive a much improved refinement of the light curves and of the derived periods (Figure 18). We performed a Lomb-Scargle (L-S: Scargle 1982) search of both archival and new light curves for periodic variable star (V427). The light curves are folded by the best-fit period of $5.4615$ and $4.4158$ days, respectively. We also calculated the false-alarm probability ($\log FAP$) for each peak and its S/N: $\log FAP_{\text{archival}} = -28.37$, $S/N_{\text{archival}} = 43.2$ for the archival data and $\log FAP_{\text{new}} = -181.07$, $S/N_{\text{new}} = 84.5$ for the new data. Since we can get better estimation with much lowered minimum FAP value, our new period is the most likely result. In the bottom panels, the resulting amplitude spectrum was calculated with PERIOD04 package. Since the archival data is more noisy than the new one, it is rather complicated to interpret the peaks of its power spectrum.

For the second example, we show the newly discovered low-amplitude pulsating variable star (Figure 19). We used the PERIOD04 package to find multiple pulsation periods. The whole process of identifying, fitting, and pre-whitening successive frequencies was repeated until no significant frequencies were found. We adopt a conservative approach in selecting the statistical significant peaks from the amplitude spectrum. A S/N amplitude ratio of 4.0 is a good criterion for independent frequencies, equivalent to 99.9% certainty of variability (Breger et al. 1993). While no clear periodicity was found in the archive data, our amplitude spectrum shows a clear excess of power centered at $22.3979 \text{days}^{-1}$ with peak amplitudes of about 1 mmag ($S/N_{\text{new}} = 9.02$).

The last example is the opposite case of the second. Figure 20 shows that this object is unlikely to be a variable source because there is no evidence for any significant peaks, which indicates that the variations are mostly noise. Extensive study on variabilities will be presented in Paper II.
6. CONCLUSION

In this paper, we introduce a new time-series photometry with multi-aperture indexing and spatio-temporal de-trending techniques, together with complex corrections to minimize instrumental biases. We used the archival, high-temporal time-series data from one-month long MMT/Megacam transit survey program. The re-calibration of the archival data has made several improvements as follows: (i) the photometric information derived from the multi-aperture indexing measurements is useful to obtain the best S/N measurement, but also to diagnose whether or not the targets are contaminated, (ii) the resulting light curves utilize nearly 100% of available data and reach precisions down to sub mmag level at the bright magnitude end without the need to throw out many outlier data points, which makes it possible to preserve data points that show intrinsic sudden variations such as flare events, (iii) corrections for position-dependent PSF variations and de-trending of spatio-temporal systematic trends improve the quality of light curves, and (iv) new photometry enables us to determine the variability nature and period estimate more accurately.

While this study deals with a particular set of data from MMT, we find our approach has a potential for other wide-field time-series observations. Multi-aperture indexing measurement is a powerful tool in isolating and even correcting various contaminations. Spatio-temporal de-trending is also very useful in removing systematics caused by PSF variation and even non-uniform extinction of thin clouds across the FOV. Chang et al. (2013) proved this for different sets of archival survey data.

This research was supported by the Basic Science Research Program of the National Research Foundation of Korea (2011-0030875). Y.-I.B. is grateful for support from the KASI-Yonsei DRC program of the Korea Research Council of Fundamental Science and Technology (DRC-12-2-KASI). We thank the MMT/M37 survey team for the kind provision of raw image data. Dr. Kim, D.-W. helped us to test the modified photometric de-trending algorithm. Additionally, we would like to thank the anonymous referee for many helpful comments, including the suggestion to use PSF photometry to further justify the successful removal of systematics (Section 3.5.2).

REFERENCES

Alard, C. 2000, A&AS, 144, 363
Alard, C., & Lupton, R. H. 1998, ApJ, 503, 325
Becker, A. C., Wittman, D. M., Boeshaar, P. C., et al. 2004, ApJ, 611, 418
Bertin, E., & Arnouts, S. 1996, A&AS, 117, 393
Bianco, F. B., Protopapas, O., McLeod, B. A., et al. 2009, AJ, 138, 568
Boulade, O., Charlot, X., Abbon, P., et al. 2003, Proc. SPIE, 4841, 72
Bramich, D. M., & Freudling, W. 2012, MNRAS, 424, 1584
