

THE ACS VIRGO CLUSTER SURVEY XVI. SELECTION PROCEDURE AND CATALOGS OF GLOBULAR CLUSTER CANDIDATES*

ANDRÉS JORDÁN^{1,2,11}, ERIC W. PENG^{3,4,5}, JOHN P. BLAKESLEE^{5,6}, PATRICK CÔTÉ⁵, SUSANA EYHERAMENDY⁷, LAURA FERRARESE⁵, SIMONA MEI⁸, JOHN L. TONRY⁹, AND MICHAEL J. WEST¹⁰

¹ Harvard-Smithsonian Center for Astrophysics, 60 Garden St., Cambridge, MA 02138, USA; ajordan@cfa.harvard.edu

² Departamento de Astronomía y Astrofísica, Pontificia Universidad Católica de Chile, Av. Vicuña Mackenna 4860, Casilla 306, Santiago 22, Chile

³ Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA

⁴ Department of Astronomy, Peking University, Beijing 100871, China

⁵ Herzberg Institute of Astrophysics, National Research Council, Victoria, BC V9E2E7, Canada

⁶ Department of Physics and Astronomy, Washington State University, 1245 Webster Hall, Pullman, WA 99163-2814, USA

⁷ Departamento de Estadística, Pontificia Universidad Católica de Chile, Av. Vicuña Mackenna 4860, Casilla 306, Correo 22, Santiago, Chile

⁸ GEPI, Observatoire de Paris, Section de Meudon, 5 place Jules Janssen, 92195 Meudon Cedex, France

⁹ Institute of Astronomy, University of Hawaii, 2680 Woodlawn Drive, Honolulu, HI 96822, USA

¹⁰ ESO, Alonso de Córdova 3107, Vitacura, Santiago, Chile

Received 2008 June 9; accepted 2008 September 3; published 2008 December 23

ABSTRACT

We present catalogs of globular cluster candidates for the 100 galaxies of the Advanced Camera for Surveys Virgo Cluster Survey, a large program to carry out imaging of early-type members of the Virgo Cluster using the Advanced Camera for Surveys (ACS) on the *Hubble Space Telescope*. We describe the procedure used to select bona fide globular cluster candidates out of the full list of detections based on model-based clustering methods with the use of expected contamination catalogs constructed using blank field observations and which are customized for each galaxy. We also present the catalogs of expected contaminants for each of our target galaxies. For each detected source we measure its position, magnitudes in the F475W (\approx Sloan g) and F850LP (\approx Sloan z) bandpasses, and half-light radii by fitting point-spread function convolved King models to the observed light distribution. These measurements are presented for 20,375 sources, of which 12,763 are likely to be globular clusters. Finally, we detail the calculation of the aperture corrections adopted for the globular cluster photometry.

Key words: catalogs – galaxies: elliptical and lenticular, cD – galaxies: star clusters – globular clusters: general – methods: statistical

Online-only material: machine-readable tables

1. INTRODUCTION

In the eleventh *Hubble Space Telescope* (*HST*) observing cycle, we initiated the Advanced Camera for Surveys (ACS) Virgo Cluster Survey (ACSVCS; Côté et al. 2004; hereafter Paper I), a program to acquire F475W (\approx SDSS g) and F850LP (\approx SDSS z) images for 100 early-type members of Virgo using the ACS (Ford et al. 1998). Paper I describes the survey itself, including a brief overview of the scientific goals, the selection of the program galaxies and their ensemble properties, the choice of filters, and the field placement and orientation.

One of the primary scientific objectives of the survey is a homogeneous study of the thousands of globular clusters (GCs) belonging to the sample galaxies. A crucial first step in this analysis is the selection of bona fide GCs from all the detected sources in a frame. This method has to be general enough to be applicable to GC systems belonging to galaxies with a wide range in properties, implying a corresponding variety in the properties of their GC systems. In particular, the size of the GC systems in the field of view varies from a few tens of GCs to thousands of them, presenting clear differences in the importance of dealing with contaminating sources such as background galaxies and foreground stars, which will, modulo cosmic scatter, be rather uniform across our sample. Beyond

keeping the contamination of the GC samples to a minimum, it is important to be able to assess, for a particular application, the expected amount of the contamination in the sample, and its expected distribution with respect to the variables being probed (e.g., size, magnitude, and color).

The data reduction procedures for the survey have been detailed in Jordán et al. (2004a; hereafter Paper II). Paper II describes the combination of the science images, the modeling and removal of the galaxy surface brightness distribution and subsequent object detection performed with SExtractor (Bertin & Arnouts 1996). An initial selection is done on the object catalogs in order to get a first set of GC candidates. These selection criteria, detailed in Section 2.6 of Paper II, are very conservative. The main aim is to remove the glaring contaminants. All the GC candidates from this first selection are subsequently run through a code (KINGPHOT) that fits PSF-convolved King (1966) models to their surface brightness profiles. This code is described in Jordán et al. (2005; Paper X). The parameters measured for each GC candidate are its magnitude, both in an aperture of $0''.2$ and the total model intensity, its celestial coordinates α and δ , its half-light radius r_h and concentration $c \equiv \log(r_c/r_t)$, where r_c and r_t are the core and tidal radii respectively. The latter is an uncertain quantity even for the brighter objects in our catalogs.

After the initial selection of GC candidates mentioned above there are still residual contaminants, mainly foreground stars and background galaxies. For the brightest galaxies, this contamination is small enough that it is not an issue, but for the faintest members of the sample it can be a significant fraction of

* Based on observations with the NASA/ESA *Hubble Space Telescope* obtained at the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5-26555.

¹¹ Clay fellow.

the overall signal; thus isolating the bona fide GC candidates becomes crucial. Traditionally, selection of GCs is accomplished by restricting in color such as to include only the expected colors of old ($\tau \approx 13$ Gyr) stellar populations with $-2.5 \lesssim [\text{Fe}/\text{H}] \lesssim 0$, and in magnitude to exclude faint, poorly measured objects. While these cuts will certainly restrict the amount of contaminants present in the samples, this procedure has some obvious drawbacks. First, unless the boundaries of the selection box are adjusted in some way or additional constraints in galactocentric distance are used, the amount of contaminants present will, in general, increase rapidly as we go from the brightest to fainter galaxies. Second, in this scheme there is no way to assess the likelihood of a given source to be a contaminant or a bona fide GC, nor in general is it easy to quantify how contamination will affect the inferred distribution functions of GC parameters.

In this work, we describe a GC selection method, which addresses the points discussed above. The method uses, in addition to a broad color cut, the measured half-light radius r_h , which proves very useful in discriminating between GCs and contaminants given the distance to our targets and the characteristics of the telescope/detector combination used. We also discuss the aperture corrections adopted for the photometry and present catalogs of GC candidates for the 100 galaxies in the ACSVCS in machine readable tables available for download from the electronic edition of the *Astrophysical Journal*. Previous papers in this series have discussed the connection between GCs and low-mass X-ray binaries (Jordán et al. 2004b; Sivakoff et al. 2007), the measurement and calibration of surface brightness fluctuations magnitudes and distances (Mei et al. 2005ab, 2007), the morphology, isophotal parameters and surface brightness profiles for early-type galaxies (Ferrarese et al. 2006a), the connection between GCs and ultracompact dwarf galaxies (Haşegan et al. 2005), the nuclei of early-type galaxies (Côté et al. 2006), the color distribution of GCs (Peng et al. 2006a), the half-light radii of GCs and their use as a distance indicator (Jordán et al. 2005), diffuse star clusters in early-type galaxies (Peng et al. 2006b), the connection between supermassive black holes and central stellar nuclei in early-type galaxies (Ferrarese et al. 2006b), and the luminosity function, color–magnitude relations, and formation efficiencies of GCs in early-type galaxies (Jordán et al. 2006, 2007; Mieske et al. 2006a; Peng et al. 2008). The GC selection method and the procedure to determine the aperture corrections described in this work have also been applied to the ACS Fornax Cluster Survey (Jordán et al. 2007b; Côté et al. 2007).

2. GC SELECTION METHOD

For each GC candidate we measure its celestial coordinates α (right ascension) and δ (declination), model magnitudes in the F475W and F850LP filters, and King model parameter estimates, half-light radii r_h and concentrations c , in each of those bands. Note that henceforth, we will use g_{475} as shorthand to refer to the F475W filter, and z_{850} denotes F850LP. Additionally, r_h will be taken to be the straight average of the g_{475} - and z_{850} -band measurements, i.e. $r_h \equiv 0.5(r_{h,z_{850}} + r_{h,g_{475}})$.

In Figure 1 we show a plot of z_{850} versus r_h for all GC candidates from the 100 galaxies in the survey after the first rough selection of Paper II. The points on this figure are culled in color by requiring $0.5 < (g_{475} - z_{850}) < 1.9$, a generous color cut that includes metallicities in the range $-2.25 < [\text{Fe}/\text{H}] < +0.56$ for all simple stellar populations

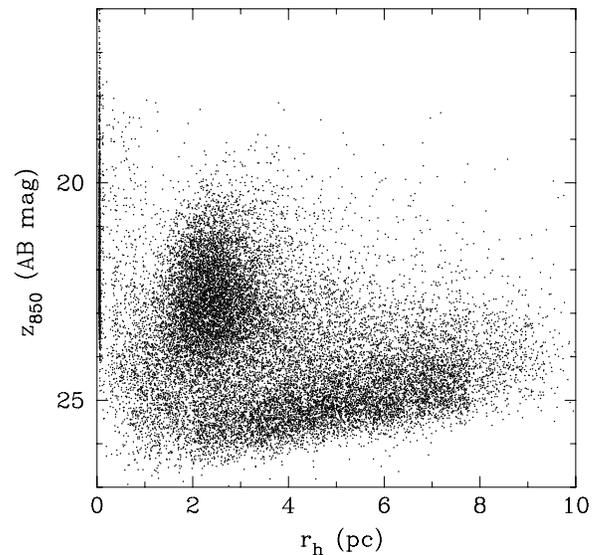


Figure 1. Distribution in the r_h – z_{850} plane of all GC candidates with $r_h < 10$ pc in the ACSVCS after the first rough selection described in Section 2.6 of Paper II. Three clusters of data points can be clearly identified: (1) a group of unresolved sources with $r_h \approx 0$, which correspond mainly to foreground stars; (2) a diagonal swath of points with faint magnitudes and large sizes, which correspond mainly to background galaxies; and (3) a group at $z_{850} \sim 20$ – 25 and $r_h \sim 3$ pc, which correspond mainly to bona fide GCs.

with ages between 2 and 13 Gyr (Bruzual & Charlot 2003), assuming either a Chabrier (2003) or Salpeter (1955) initial mass function. We also cull in r_h by requiring $r_h < 10$ pc, a cut that is also very inclusive for typical GCs (see Jordán et al. 2005). Note that although we will quote values of r_h in pc throughout this work for convenience, we are really always working with angular measurements. For the purposes of this work, angular measurements are transformed to pc adopting a distance of 16 Mpc.¹² The conversion factor from pc to arcsec is then $78 \text{ pc arcsec}^{-1}$. While in principle we could use the measured surface brightness fluctuations distances from our survey (Mei et al. 2007) to convert angular to physical distances for each galaxy individually (as we have done in many of the survey papers), we have chosen not to do so for GC classification. The reason is that one of our scientific objectives is to test some characteristics of the GC systems (e.g., the GC luminosity function) as distance indicators and thus we decided it was best not to use any distance information in the selection of GC candidates to avoid possible biases in our results.

Three distinct groupings of data points can clearly be distinguished in Figure 1. At $r_h \sim 0$ there is a vertical clustering of points which corresponds to unresolved sources, i.e., sources that are most likely foreground stars. At $r_h \sim 3$ pc and $z_{850} \sim 20$ – 25 mag there is a second group of points which corresponds to the GC population of the Virgo galaxies. Lastly, there is a diagonal swath of points which corresponds to faint, background galaxies. The diagonal shape of the faint end envelope of this group is due to a completeness effect in surface brightness, as at a given magnitude more extended sources are less likely to be detected. Figure 1 clearly shows the usefulness of having size information in order to separate GCs from background galaxies and foreground stars, as the different types of objects separate into three distinct groupings. We note that, in this context, the

¹² In Mei et al. (2005a, 2007) we adopt a mean distance of 16.5 Mpc to Virgo. The work presented here was performed before we adopted this value, but it is well within its systematic uncertainties. The effect in the classification of sources of adopting 16.5 Mpc instead of 16 would be in any case very small.

use of z_{850} -band magnitudes is preferred to that of g_{475} band due to the fact that background galaxies are in general bluer than typical GCs, and thus they are fainter with respect to GCs in z_{850} than in g_{475} .

It is evident from Figure 1 that the unresolved sources can easily be eliminated by requiring $r_h > 0.75$ pc, which we adopt as the minimum angular size to be considered a bona fide GC. This is close to the angular limit down to which we can reliably measure r_h for GCs (Jordán et al. 2005). This cut leaves the task of separating the GCs from the background galaxies. For this, we use the clustering method described below.

2.1. Clustering Method

After removing the data cluster corresponding to unresolved sources, the data are a mixture of points drawn mainly from two populations, namely the GCs and a population comprised mostly of background galaxies, which we will henceforth term “contaminants.” Thus, their joint distribution can be modeled using a mixture model (see below) with two components, in which the total observed distribution in the r_h - z_{850} plane is the result of summing these two components weighted by their respective sizes. In what follows we briefly outline some formalism regarding mixture models in general and the specialization to our problem in particular. The discussion that follows draws on Fraley & Raftery (2002), to which the reader is directed for more details.

In general, we can model a random variable \vec{Y} as a mixture of N components

$$\begin{aligned} \vec{Y}_1 &\sim d_1(\theta_1) \\ \vec{Y}_2 &\sim d_2(\theta_2) \\ &\vdots \\ \vec{Y}_N &\sim d_N(\theta_N) \\ \vec{Y} &= z_1 \vec{Y}_1 + \dots + z_N \vec{Y}_N \end{aligned}$$

where $(z_1, \dots, z_N) \in \{0, 1\}$ indicate the component membership of \vec{Y} ($z_i = 1$, for only one i), and d_k and θ_k are the probability densities and parameters of the k th component, respectively. The z_i are assumed to be distributed as a multinomial of one draw from the N components with $Pr(z_1 = 1) = f_1, \dots, Pr(z_N = 1) = f_N$, where f_k is the probability that a given observation of \vec{Y} belongs to the k th component ($f_k > 0$ and $\sum_k f_k = 1$).

In order to estimate the parameters of a mixture model $(\theta_1, \dots, \theta_N, f_1, \dots, f_N)$ via maximum likelihood, the expectation-maximization (EM) algorithm is employed (Dempster et al. 1977; McLachlan & Krishnan 1997). This algorithm, which alternates between two steps, an “E” (expectation) and an “M” (maximization) step, is a general approach to maximum likelihood maximization in which the data consist of n observations \vec{x}_i that arise from (\vec{y}_i, \vec{z}_i) , in which the \vec{y}_i are observed and the \vec{z}_i are not observed. In our case, none of our observables tell us directly to which of the groups a given data point belongs, but in our statistical description there will be two groups, and a given data point will belong to one of them. Thus, we have that $z_{ik} = 1$ if \vec{y}_i belongs to group k , and 0 otherwise.

Having an indicator variable that gives information on group membership is necessary in order to estimate the parameters θ_k of a given component. Take for instance the simple example of estimating the mean for a multivariate normal: it is clear that in calculating the mean for group k , only data points belonging to that group have to be considered, i.e. $\mu_k = (1/n_k) \sum_i z_{ik} \vec{y}_i$, where $n_k \equiv \sum_i z_{ik}$.

Under quite general conditions, which are satisfied in our case, the *observed* data likelihood $\mathcal{L}_O(\vec{y}|\theta)$ can be obtained from the *complete* data likelihood $\mathcal{L}_C(\vec{x}|\theta)$ as $\mathcal{L}_O(\vec{y}|\theta) = \int \mathcal{L}_C(\vec{x}|\theta) d^k z$; the maximum likelihood estimate of θ based on the observed data maximizes \mathcal{L}_O . For a sample of n independent multivariate observations $\vec{y}_1, \dots, \vec{y}_n$ the observed likelihood of a mixture model with N components is

$$\mathcal{L}(\theta_1, \dots, \theta_N; f_1, \dots, f_n | \vec{y}) = \prod_{i=1}^n \sum_{k=1}^N f_k d_k(\vec{y}_i | \theta_k). \quad (1)$$

The corresponding log likelihood of the *complete* data is

$$l(\theta_k, f_k, z_{ik} | \vec{x}) = \sum_{i=1}^n \sum_{k=1}^N z_{ik} \ln(f_k d_k(\vec{y}_i | \theta_k)). \quad (2)$$

This log likelihood is the one maximized by the EM algorithm, which can be shown to converge to local maximum of the observed data likelihood under mild regularity conditions. The E, or expectation step, is given by the assignment

$$\hat{z}_{ik} \leftarrow \frac{\hat{f}_k d_k(\vec{y}_i | \hat{\theta}_k)}{\sum_{j=1}^N \hat{f}_j d_j(\vec{y}_i | \hat{\theta}_j)}, \quad (3)$$

where a hat indicates as usual in statistical work an estimate of a parameter. \hat{z}_{ik} is the conditional expectation value of z_{ik} , i.e. an estimate given all other parameters of the group membership of \vec{y}_i . The M, or maximization, step corresponds to maximizing the expression given in Equation (2) over f_k and θ_k with the z_{ik} fixed at the values \hat{z}_{ik} obtained in the E step. The E and M steps are then iterated until the parameters have converged. The value of \hat{z}_{ik} at a maximum of Equation (1) is the estimated probability that observation i belongs to group k . This quantity can then be used to classify a given observation \vec{y} into its most likely group.

It is a common practice to choose the d_k to be multivariate normals ϕ_k with mean $\vec{\mu}_k$ and covariance matrix Σ_k . Explicitly,

$$\begin{aligned} \phi_k(\vec{y} | \vec{\mu}_k, \Sigma_k) &= (2\pi)^{-p/2} \det(\Sigma_k)^{-1/2} \\ &\times \exp\left(-\frac{1}{2}(\vec{y} - \vec{\mu}_k) \Sigma_k^{-1} (\vec{y} - \vec{\mu}_k)\right). \end{aligned} \quad (4)$$

That a description using two multivariate normal components would be rather reasonable in our case can be seen in Figure 2 where a two-dimensional kernel density estimate of the sample with the unresolved sources already removed is shown.

2.1.1. Specialization to GC Classification in the ACSVCS

For our problem we have two components, i.e. $k = 1, 2$, with the first component being the GCs and the second the contaminants. Given models for the distributions d_1 and d_2 for each component, we can characterize them by the set of parameters $\theta_k = (\vec{\mu}_k, \vec{s}_k)$, $k = 1, 2$, where $\vec{\mu}_k = (\mu_{r_h, k}, \mu_{z_{850}, k})$ are location (mean) parameters characterizing the distribution in the r_h - z_{850} plane and \vec{s}_k are any set of shape parameters characterizing the distributions.

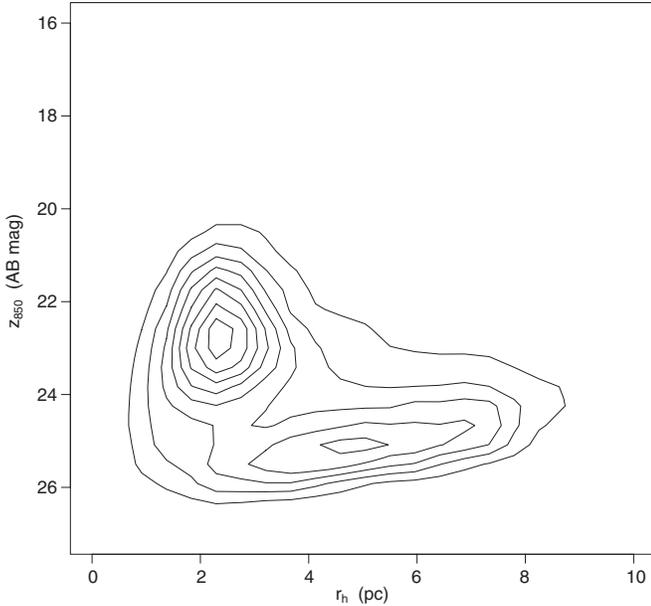


Figure 2. Contour plot of a two-dimensional kernel density estimate of the distribution of points in the r_h - z_{850} plane after removal of the unresolved sources. Two distinct components can be seen whose shapes suggest that using a multivariate normal mixture model can provide a meaningful representation of the total distribution.

We will assume that the shape parameters \vec{s}_k of the GC distribution are universal, i.e. we assume the rough *shape* of the clusters of GC points in the r_h - z_{850} plane is independent of the galaxy. This is a reasonable assumption, and a necessary one as well, as the faint galaxies do not have in general enough signal in the data cluster corresponding to GCs in order to reliably determine the form of its joint r_h and z_{850} distribution.

While the *shape* of the distributions will be kept constant, it might be useful to let some components of $\vec{\mu}_k$ depend on the particular galaxy under study. First, the mean magnitude of GCs will obviously depend on the distance and on the typical mass and metallicity of the GCs. Also, we cannot assume that the mean GC r_h will be the same in all galaxies (indeed they are not; see Jordán et al. 2005). In the case of the contamination group, the mean magnitude and size may vary because of the varying levels of the galaxy surface brightness: the contamination group in a dwarf galaxy should have a fainter mean z_{850} due to the decreased mean surface brightness in the frame (Ferrarese et al. 2006a). In practice, we found that letting $\mu_{z_{850},1}$ to be a free parameter leads to unsatisfactory solutions in some cases, and we therefore chose to fix it at $\mu_{z_{850},1} = 22.7$ mag for all galaxies, but left $\mu_{r_h,1}$ to be a free parameter. When letting $\mu_{z_{850},1}$ be free, the dispersion around the assumed value of 22.7 mag was significantly smaller than the assumed dispersion $\sigma = 1.3$ mag of the magnitude distribution of GCs (see below), and therefore the net effect on the final GC classification is small.¹³

While modeling the two clusters with multivariate normals might provide a good first approximation as suggested by Figure 2, there is the concern that the steep decline of a multivariate Gaussian in the r_h direction might result in the incorrect classification of points at large r_h as contamination. Indeed, the dis-

tribution of r_h for the Milky Way shows an extended tail toward large r_h (see e.g., Jordán et al. 2005).

Instead of relying on multivariate normals to describe the GC distribution, we can instead make use of important prior information. It is well known that the luminosity function of GCs is rather universal, being reasonably well approximated by a Gaussian with $\sigma \sim 1.3$ mag (Harris 2001).¹⁴ Additionally, it is known that the half-mass radii of GCs are uncorrelated with their luminosity for $M \lesssim 2 \times 10^6 M_\odot$, a mass limit which includes the great majority of GCs (McLaughlin 2000; Jordán et al. 2005).¹⁵ Therefore, magnitude and r_h can be taken as independent, and the joint distribution can be taken as the product of a magnitude distribution and an r_h distribution. Given these points, we have adopted the following model for the distribution function of the GCs, d_{GC} ,

$$d_{GC}(z_{850}, r_h | \mu_z, \mu_{r_h}) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(z_{850} - \mu_z)^2}{2\sigma^2}\right] g_{gc}(r_h), \quad (5)$$

where g_{gc} is the distribution function of r_h and $\sigma = 1.3$ mag. In order to model g_{gc} we determine it empirically from our data by taking all GC candidates in M87 (VCC 1316) and M49 (VCC 1226) satisfying $z_{850} < 23$ mag. This sample is composed almost exclusively of GCs, contamination being almost negligible for these two galaxies. Using this sample, we determined a nonparametric form for g_{gc} using a normal kernel density estimate (Silverman 1986). Even though the combined GC candidate sample of these two giant galaxies has negligible contamination as a whole, for $r_h \gtrsim 6$ pc contamination is potentially an issue. Thus, we used the nonparametric density estimate for $r_h < 6$ pc only. For larger r_h , we extended the distribution with a power-law of the form r_h^{-p} , the parameter p determined by fitting to the observed distribution of half-light radii with $4 < r_h < 6$ pc. The final g_{rh} we used is shown in Figure 3. The form of this distribution and the power law behavior for large GCs ($r_h > 4$ pc) is consistent with the parametric form presented by Jordán et al. (2005) for the size distribution of GCs. The tail of this distribution is important to classify correctly extended sources in our procedure.

For the contaminants, we model their distributions using customized control fields. The pipeline described in Paper II was run on a series of control fields, which are listed in Table 1 of Peng et al. (2006a; Paper IX). The source-detection algorithm was run in these 17 fields using the weight maps W'_{ij} constructed for each of the survey galaxies in turn. The weight maps, whose construction is detailed in Paper II, encode the detection thresholds of each galaxy and therefore we obtain for each of the galaxies a catalog of contaminants that we would have observed *had the galaxy been present in the control fields*. In this sense, we are able to customize the control fields for each of our galaxies. Using the catalog of expected contaminants for each galaxy, we build a two-dimensional kernel density estimate d_{cont} to represent the joint distribution function of the contaminants. This distribution has no free parameters and is thus kept fixed during the source classification process for each galaxy.

¹³ $\lesssim 0.5\%$ of the sources would switch classification when performing the classification by letting $\mu_{z_{850},1}$ free instead of being fixed.

¹⁴ We have shown using the ACSVCS GC catalogs based on the classification algorithm described in this work that there is actually a clear σ - M_B relation, in the sense that fainter galaxies have narrower GC luminosity functions (Jordán et al. 2006, 2007a).

¹⁵ For objects with higher masses, which are now termed ultracompact dwarfs or dwarf-globular transition objects, it is found that a r_h - L relation emerges (Haşegan et al. 2005; Mieske et al. 2006b, Kissler-Patig et al. 2006).

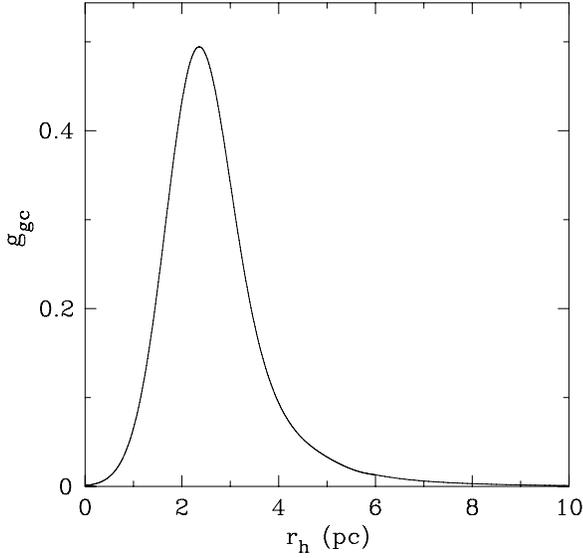


Figure 3. Adopted probability density function g_{gc} of the GC half-light radii. This distribution was constructed using a kernel density estimate for GC candidates from VCC 1226 (M49) and VCC 1316 (M87) for $r_h < 6$ pc. Above that angular radius we extended the distribution with a power-law r_h^{-p} whose index $p = -5$ was determined by fitting to the kernel density estimate of the distribution in the range $4 < r_h < 6$ pc. The form of g_{gc} is consistent with the parametric form presented for this distribution in Jordán et al. (2005).

With the form of the density functions for each galaxy in hand, we need to estimate the parameters of the mixture model. The procedure we adopted is the following. First, the whole GC candidate sample of all galaxies (minus the “stars”) is used to determine a mixture model of the form of Equation (1) using multivariate normals and two groups (GCs and contaminants). This model is constructed in order to then provide a two-component, zero-order, mixture model that will be used to provide (given the $z_{850}-r_h$ data points of any given galaxy) initial values for the unobserved group indicator variables z_{ik} . In order to estimate the means $\vec{\mu}_k$, $k = 1, 2$ and covariance matrices Σ_k , $k = 1, 2$ we separated the data points into two groups using a hierarchical model-based clustering method (Fraley 1998). The means of the multivariate mixture model thus obtained are $\vec{\mu}_1 = (2.47, 22.93)$ and $\vec{\mu}_2 = (5.27, 24.66)$, where 1 denotes the GC group and 2 the contaminants. The covariance matrices are given by

$$\Sigma_1 = \begin{pmatrix} 1.257 & 0.192 \\ 0.192 & 0.637 \end{pmatrix} \quad \text{and} \quad \Sigma_2 = \begin{pmatrix} 0.409 & 0.420 \\ 0.420 & 1.823 \end{pmatrix}. \quad (6)$$

The resulting separation into two groups predicted by the multivariate mixture model applied to the whole sample is shown in Figure 4.

At this point we have a multivariate mixture model of the form given in Equation (1) in hand that given a set of points in the size-magnitude plane will classify them into either group according to the estimates of z_{ik} given by Equation (3). For each galaxy we use this model to provide estimates for the z_{ik} that in turn provide the initial conditions for the M step in the EM algorithm that is used to estimate the parameters of the final mixture model for that galaxy (using the distributions d_{GC} and d_{cont}). In other words, the multivariate Gaussian model estimated using the whole sample is used to provide the initial conditions that are needed in order to maximize the likelihood given by (see Equation (2))

Table 1
Maximum Likelihood Parameters of Mixture Model

VCC (1)	μ_{r_h} (pc) (2)	f_{GC} (3)	VCC (1)	μ_{r_h} (pc) (2)	f_{GC} (3)
1025	2.24	0.676	1743	2.75	0.244
1030	2.68	0.595	1779	3.45	0.165
1049	3.69	0.267	1826	3.46	0.231
1062	3.09	0.622	1828	3.46	0.307
1075	3.48	0.346	1833	3.44	0.281
1087	3.47	0.452	1857	3.48	0.212
1125	3.31	0.479	1861	3.48	0.471
1146	2.85	0.572	1871	2.77	0.258
1154	2.77	0.680	1883	3.23	0.459
1178	3.03	0.524	1886	3.51	0.223
1185	3.45	0.317	1895	3.32	0.208
1192	3.23	0.670	1903	2.79	0.770
1199	3.08	0.666	1910	3.33	0.435
1226	2.72	0.941	1913	3.44	0.497
1231	2.85	0.763	1938	3.12	0.552
1242	3.08	0.588	1948	3.48	0.215
1250	3.30	0.447	1978	2.56	0.949
1261	2.98	0.394	1993	2.97	0.116
1279	2.82	0.664	200	3.41	0.306
1283	2.76	0.494	2000	3.01	0.707
1297	3.41	0.699	2019	3.48	0.386
1303	2.84	0.437	2048	3.44	0.295
1316	2.73	0.971	2050	3.53	0.284
1321	3.09	0.421	2092	2.91	0.497
1327	3.42	0.704	2095	3.09	0.509
1355	3.59	0.314	21	3.50	0.274
140	3.46	0.305	230	3.34	0.386
1407	3.28	0.479	33	3.45	0.186
1422	3.68	0.335	355	2.77	0.401
1431	3.27	0.495	369	3.10	0.639
1440	3.30	0.357	437	3.54	0.478
1475	2.82	0.525	538	3.15	0.178
1488	3.51	0.249	543	3.47	0.285
1489	3.54	0.276	571	2.98	0.208
1499	3.60	0.324	575	2.54	0.344
1512	2.37	0.211	654	3.08	0.390
1528	3.43	0.426	685	2.94	0.701
1535	2.98	0.585	698	3.25	0.600
1537	2.66	0.480	731	2.23	0.922
1539	3.48	0.494	751	3.03	0.232
1545	3.39	0.511	759	2.86	0.574
1619	2.85	0.464	763	2.58	0.900
1627	2.91	0.226	778	2.62	0.533
1630	2.46	0.487	784	2.87	0.529
1632	2.62	0.851	798	3.04	0.605
1661	3.46	0.264	828	2.99	0.542
1664	2.90	0.666	856	3.50	0.421
1692	2.90	0.657	881	2.77	0.792
1695	3.44	0.235	9	3.44	0.248
1720	3.27	0.427	944	2.77	0.568

Notes.

Key to columns: (1) Galaxy VCC number; (2) Mean half-light radius of GC component (assuming $D = 16$ Mpc); (3) Estimated fraction of the total sample of the GC component.

The corresponding quantity for the contaminants component, f_{cont} is given by $f_{cont} \equiv 1 - f_{GC}$.

$$\mathcal{L}(\vec{\mu}_{GC}; f_{GC}, f_{cont} | \vec{y}) = \prod_{i=1}^n [p_{i,GC} f_{GC} d_{GC}(\vec{y}_i | \vec{\mu}_{GC}, \sigma) + p_{i,cont} f_{cont} d_{cont}(\vec{y}_i)] \quad (7)$$

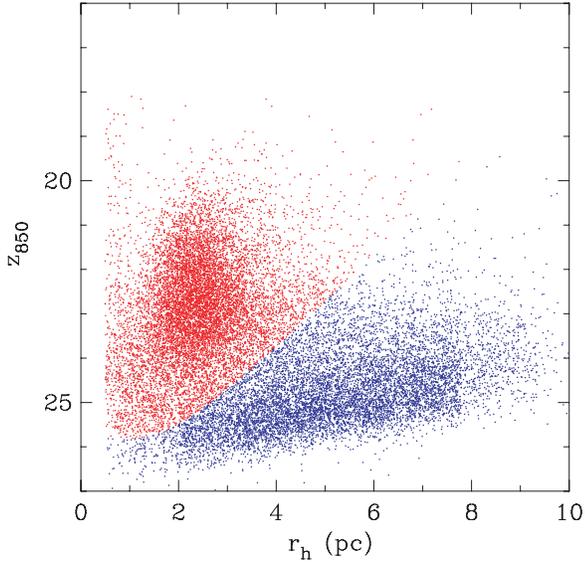


Figure 4. Distribution in the r_h - z_{850} plane of all GC candidates with $r_h < 10$ pc in the VCS after the first rough selection described in Section 2.6 of Paper II, and with unresolved sources ($r_h < 0.75$ pc) removed. Using a multivariate normal mixture model the sample is divided into two groups: GCs (red points) and “contaminants” (blue points; mainly background galaxies). This is used to define the initial mixture model for each galaxy.

and thus obtain estimates for f_{GC} , f_{cont} , $\mu_{r_h,GC}$ using the EM algorithm. Note again that d_{cont} has no free parameters and that $p_{i,cont} \equiv (1 - p_{i,GC})$. The estimates we obtain for these parameters are listed in Table 1 for each of our program galaxies.

After the parameters $\mu_{r_h,GC}$, μ_{cont} , f_{GC} , and f_{cont} have been estimated we assign for each point $\vec{y} = (r_h, z_{850})$ in that galaxy’s sample the probability of it being a GC given by (see Equation (3))

$$p_{GC} = \frac{f_{GC}d_{GC}(\vec{y} | \vec{\mu}_{GC}, \sigma)}{f_{GC}d_{GC}(\vec{y} | \vec{\mu}_{GC}, \sigma) + f_{cont}d_{cont}(\vec{y})}, \quad (8)$$

and, given that there are just two components, the corresponding probability p_{cont} of it being a contaminant is given simply by $p_{cont} = 1 - p_{GC}$.

A final step in the classification is that $p_{GC} \equiv 1$ is assigned to all sources satisfying $z < 23$ mag and $1.5 \text{ pc} < r_h < 4$ pc, as we want to consider these sources to be bona fide GC candidates regardless of the exact value of p_{GC} returned by the algorithm (we note that of the 6475 sources satisfying these conditions only five originally have $p_{GC} < 0.5$). Due to the high level of contamination of faint extended objects we also set $p_{GC} = 0$ for $z_{850} > 25.15$ mag, $g_{475} > 26.35$, and $r_h > 10$ pc. Sources fainter than that magnitude limit are almost certainly contamination, and in any case contain little useful information due to the large errors in their measured quantities, while sources more extended than the size limit are hard to select against the majority of background galaxies with those angular sizes (see Figure 5).

What are the advantages of this method over just defining fixed boundaries in the r_h - z_{850} plane? By estimating f_{GC}

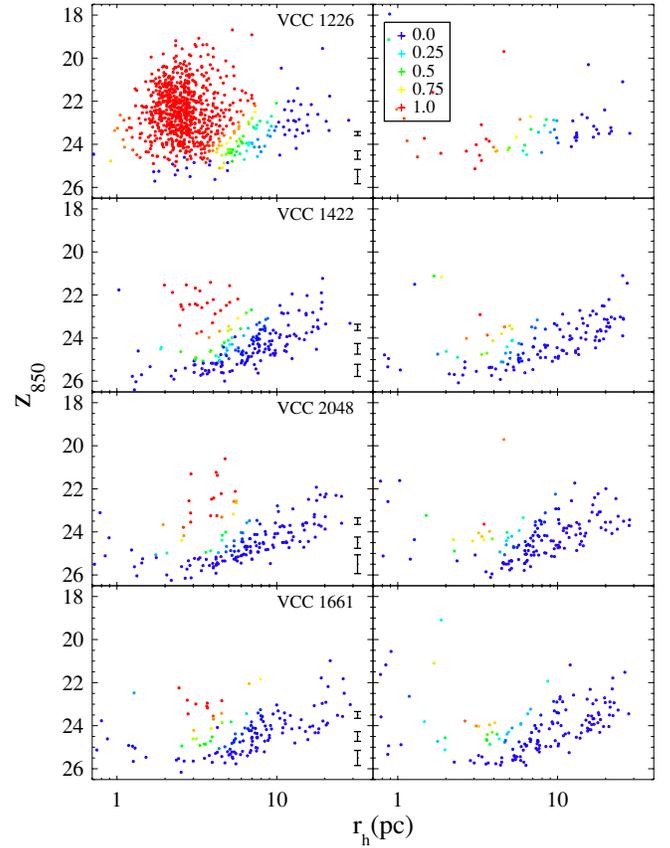


Figure 5. Size magnitude diagrams for four representative galaxies of the ACSVCS. Using the maximum-likelihood estimates of the parameters of a mixture model as described in the text we assign a probability p_{GC} for each source that it is a GC. Objects are color coded according to this probability. On the left are the objects in our program fields, and on the right are the objects in the custom control fields for that galaxy, scaled to a single field. Note how the customization of the control fields for each galaxy is able to reproduce the different shapes of the contamination component due to the varying levels of mean underlying surface brightness.

and f_{cont} for each galaxy, we naturally include the fact of varying ratios of contamination to GC candidates. In addition, by letting the $\mu_{r_h,GC}$ be determined for each galaxy we naturally account for variations in the mean r_h as well. Also, we have a *quantitative measure* of how likely a certain data point is to be a GC (under the assumed model) and this knowledge can be included in statistical estimators applied to the GC sample.

In order to illustrate the performance of the method we show the results for four galaxies that span the magnitude range of our sample: VCC 1226 (= M49), VCC 1422, VCC 2048 and VCC 1661. In terms of apparent blue luminosity, these are the 1st, 50th, 51th and 100th ranked galaxies in the ACSVCS sample, respectively. In Figure 5 we show the resulting classification in the r_h - z_{850} plane for these galaxies along with the same classification applied to our custom control fields (scaled to a single field). In Figure 6, we show the resulting GC luminosity functions and color distributions when restricting the objects to those having $p_{GC} \geq 0.5$. In the samples thus selected it will still be necessary to take into account the residual contamination that is classified as bona fide GCs (i.e., the false positives), but it should be clear from these figures that this contamination has been greatly reduced by our selection using a process that naturally takes into account the richness of the GC system of each galaxy.

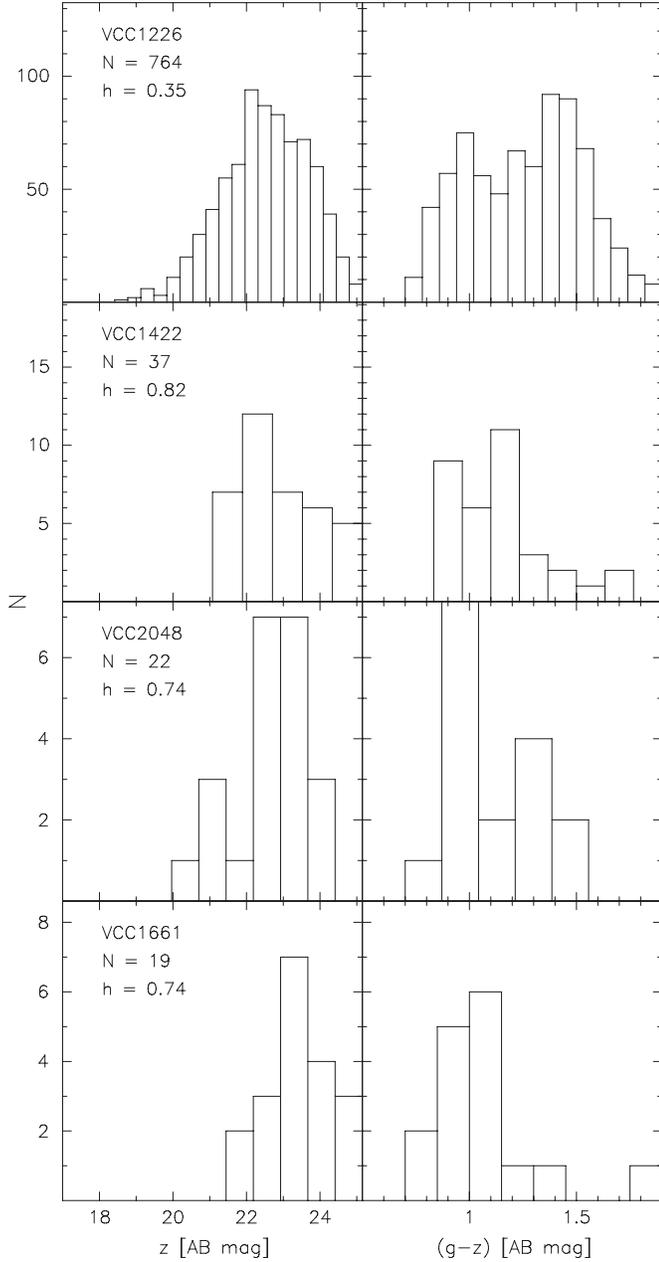


Figure 6. GC luminosity functions (left panels) and color distributions (right panels) for all GC candidates satisfying $p_{GC} \geq 0.5$ in galaxies spanning the luminosity range of our sample. The galaxy names are indicated in the left panels, where we also indicate the total number of objects N and the binwidth h used to construct the histograms.

While the selection method effectively isolates the desired GC data cluster, studies that aim to study the shape of the GC distributions in either size or magnitude need to test their conclusions against any subtle biases that the selection of GCs might have imposed on them via the choice of the form of d_{GC} . This can be easily done by considering the robustness of results when selecting alternate GC samples of GCs that do not rely on p_{GC} , as we have done when studying the luminosity function of GCs (Jordán et al. 2006, 2007a).

3. GC PHOTOMETRY: APERTURE CORRECTIONS

The photometric zeropoints and foreground reddening corrections that we adopted are detailed in Sections 2.6 and 2.7 of

Jordán et al. (2004a). A challenging aspect of obtaining accurate photometry of marginally resolved objects, such as GCs at the distance of the Virgo cluster, is that their aperture corrections depend on their intrinsic size and luminosity profile, which are not known a priori for each source. The usual practice in previous photometric work on GCs with *HST* has been to adopt an average correction obtained from bright, high signal-to-noise ratio GC candidates or, alternatively, from simulated King models with $r_h \approx 3$ pc and assumed fixed concentration. In the ACSVCS, we provide two measurements of the magnitude per filter for each GC candidate. The first measurement is of the total magnitude using the best-fit King (1966) model, a method which takes into account the proper aperture correction for each individual object. We refer to these magnitudes as “model magnitudes.” The second measurement is of the best-fit aperture magnitude within a 4 pixel ($= 0''.2$) radius, with an aperture correction appropriate for an average GC ($r_h = 3$ pc, $c = 1.5$). We refer to these as “average correction aperture magnitudes.” The former method is best for obtaining an unbiased total GC flux. The latter method is useful because it relies only minimally on the King model fitting process, and thus may be better for studies concerned with the colors of faint GCs. We outline both methods below, and discuss the differences in their use.

3.1. Model Magnitudes: Total Magnitudes from King Model Fits

We obtain model magnitudes using the best-fit King model for each GC candidate in conjunction with the appropriate, size-dependent aperture corrections. As described in the Appendix of Jordán et al. (2005), the fitted King models are convolved with the appropriate point-spread function (PSF). These PSFs are constructed in the manner outlined by Jordán et al. (2004a, Section 2.6), and extend to a radius $r = 0''.5$. We fit the observed light distributions of GC candidates within a fitting radius r_{fit} (see the Appendix of Jordán et al. 2005), which means that in a strict sense, the PSF-convolved model matches the amount of light within that radius, with the rest (out to $r = 0''.5$) added as prescribed by the PSF-convolved model. *These magnitudes already take into account the size (r_h) of each object.* However, aperture corrections are still needed for the following reasons. First, the magnitudes require the aperture correction of the PSF from $r = 0''.5$ to infinity. Second, the PSF we have used in the fitting will have some differences from the one used to define the aperture corrections in Sirianni et al. (2005), which we assume to be a representation of the “true” mean PSF.

To obtain the aperture corrections arising from these effects, we adopted the following procedure. We constructed a set of King models of various r_h and convolved them with the PSF $\mathcal{P}(\vec{x})$ that is used for fitting the GC candidates in each of the z_{850} and g_{475} bands. Then, we convolved the same set of models with a PSF $\mathcal{S}(\vec{x})$ that was constructed up to a radius of $3''$ using stars in the outskirts of 47 Tucanae (as used in Mei et al. 2005). Aperture corrections derived from these PSFs (\mathcal{S}) are consistent with those derived by the PSFs used in Sirianni et al. (2005). Given a fitting radius r_{fit} , the aperture correction \mathcal{A} appropriate for a GC candidate of half-light radius r_h and concentration c described by a King model $k(\vec{x} | r_h, c)$ is given by

$$\mathcal{A}(r_h, c, r_{fit}) = -2.5 \log_{10} \left[\frac{\int_{|\vec{x}| < r_{fit}} [k(\cdot | r_h, c) \otimes \mathcal{P}](\vec{x}) d^2x}{\int_{|\vec{x}| < r_{fit}} [k(\cdot | r_h, c) \otimes \mathcal{S}](\vec{x}) d^2x} \right], \quad (9)$$

where \otimes denotes convolution. All King models k and PSFs

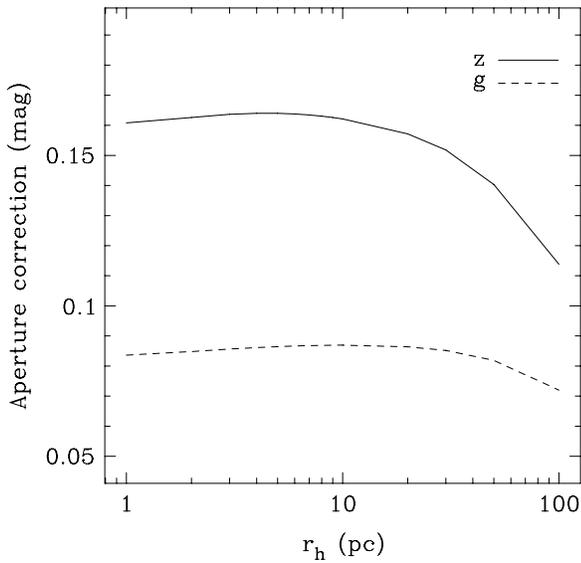


Figure 7. Aperture corrections to produce model magnitudes, calculated according to Equation (9) for a fitting radius of 4 pixels as a function of the object’s half-light radius r_h . The solid line shows the corrections for the z_{850} band while the dashed line shows the correction for the g_{475} band.

(\mathcal{P} , \mathcal{S}) are normalized to have a total flux of unity. Thus, the numerator within the logarithm in Equation (9) is the fraction of the PSF(\mathcal{P})-convolved model flux that is within the fitting radius. Correspondingly, the denominator is the fraction of the PSF(\mathcal{S})-convolved model flux that is within the fitting radius. The ratio of the two represents the aperture correction necessary to transform the fitted magnitude to a total model magnitude. Remember that the fitted magnitude that we are correcting assumes the PSF, \mathcal{P} , and already incorporates the numerator term. We define aperture corrections as values to be subtracted from the fitted magnitude.

In order to apply this correction for our GC candidates we assumed a fixed concentration $c = 1.5$ for all GCs, and then computed \mathcal{A} for $r_h = 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 50, 100$ pc and $r_{\text{fit}} = 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15$ pixel and interpolated from these values to apply an aperture correction for any given cluster with best-fit r_h fitted using a radius r_{fit} .

The resulting aperture corrections for both bands are illustrated in Figure 7. There is a very mild dependence on the aperture correction on the measured r_h for GCs (which are all required to have $r_h < 10$ pc), but there is some change of the order of a few hundredths of a magnitude for very extended objects. Thus, our aperture corrections are essentially equivalent to having applied an average correction for all clusters. Note that the latter is true only because the fitted magnitudes include the amount of light outside the fitting radius (within $r < 0''.5$) as given by the best fit PSF-convolved King model. It would not be true if an average aperture correction is applied to a simple aperture magnitude, as has usually been done in previous photometric measurements of GC systems at the distance of Virgo. In that case, the aperture correction would get systematically larger as r_h increases. We note that, as expected, at $r_h \sim 0$ the aperture corrections are similar to those expected for point sources from $0''.5$ to infinity (Siriani et al. 2005). We note also that assuming a fixed c when deriving our aperture corrections is not a strong assumption: for $r_h < 10$ pc, aperture corrections vary by less than 0.005 mag when varying c between 1 and 2.

3.2. Average Correction Aperture Magnitudes

In addition to measuring total magnitudes based on the best-fit PSF convolved King models, we also measured the best-fit amount of light in a 4 pixel ($= 0''.2$) aperture radius, and then applied an aperture correction appropriate for an *average* object. We note that the light measured within the aperture is measured on the best-fit King model rather than the data. This method has the advantage of dealing naturally with bad pixels and subpixel shifts that can affect aperture magnitudes measured directly from the data. The normalization of the model (i.e., the total flux within the fitting radius) is the most robustly fitted quantity even for our faintest objects, and because the 4 pixel aperture is almost always equal in size to the fitting radius (sometimes smaller), the aperture flux measured on the model is very reliable. These magnitudes are best used for studying the colors of faint GCs, as they provide a color measurement that relies only weakly on the fitting process, and thus can have smaller photometric errors compared to colors derived from model magnitudes.

To estimate the total magnitudes from these aperture magnitudes, we calculate the aperture correction for a King model with $r_h = 3$ pc and $c = 1.5$, which is the quantity \mathcal{A}_{ap} given by

$$\mathcal{A}_{\text{ap}} = -2.5 \log_{10} \left[\int_{|\bar{x}| < 0''.2} [k(\cdot | 3 \text{ pc}, 1.5) \otimes \mathcal{S}](\bar{x}) d^2x \right]. \quad (10)$$

The values obtained for \mathcal{A}_{ap} are then 0.237 mag for the g_{475} band and 0.347 mag for the z_{850} band. Assuming $c = 1$ or $c = 2$ would only make a minimal difference (< 0.005 mag). We note that model magnitudes and average corrected aperture magnitudes are consistent in the average as expected.

3.3. Further Notes on Aperture Corrections

As we have discussed, photometry of marginally resolved sources whose individual sizes are not known a priori is a challenge for *HST* studies of GCs at the distance of the Virgo Cluster. In principle, standard aperture photometry, where one measures the flux in a given aperture and applies the same aperture correction for all objects, is subject to biases such that the total fluxes of extended objects are underestimated. We emphasize that *our model magnitudes fully account for this effect*, by fitting for the size of each individual object, and applying the proper aperture corrections out to a radius of $3''$ using a suite of PSF-convolved King models. Studies concerned with the magnitudes or colors of sources that are large ($r_h \gg 3$ pc), and where size-dependent biases may be important, should use model magnitudes. For example, Mieske et al. (2006a), the ACSVCS study of the color–magnitude relations in GC systems, solely used model magnitudes to avoid any size-dependent biases on the magnitudes or colors.

Our average corrected aperture magnitudes, however, are more similar to previous work, where the flux is measured in a single aperture and an average aperture correction is applied. This presents the possibility that if GCs have sizes significantly different from $r_h = 3$ pc, the size at which we calculate our fiducial correction, then their magnitudes and colors could be systematically in error. The left panel in Figure 8 shows the aperture corrections in both filters to a 4 pixel aperture for King models with a range of sizes. This plot illustrates how GCs with sizes substantially larger than 3 pc can have significantly underestimated fluxes as measured by average corrected aperture magnitudes (~ 1 mag for $r_h = 20$ pc). *This is*

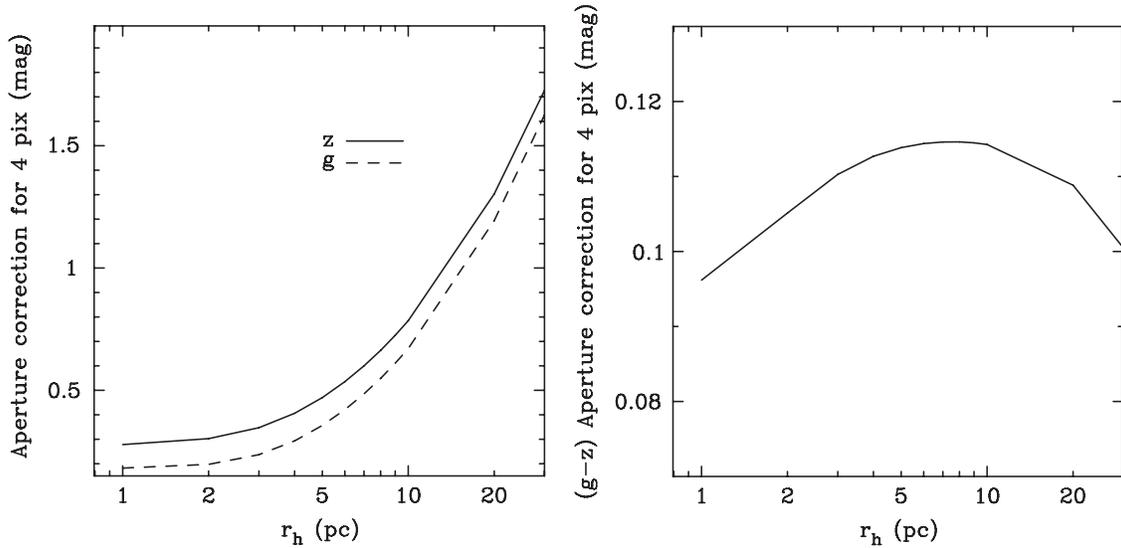


Figure 8. Left: aperture correction that would be required if measuring the flux in an aperture of 4 pixels for PSF-convolved King models of different r_h and $c = 1.5$. Our model magnitudes already take into account the light outside the fitting radius and as a consequence they require an aperture correction that varies very mildly with r_h (see Figure 7). Right: required aperture corrections for the $g-z$ color when calculating the color from magnitudes measured in a 4 pixel aperture, as a function of the object’s half-light radius r_h . Even though the individual magnitudes have aperture corrections that depend strongly on the object’s size, aperture-based colors show variations at the < 0.01 mag level only.

why in all our ACSVCS studies, we always use model magnitudes for measures of total flux. The bias in color is the difference of the two curves in the left panel of Figure 8, and we show this in the right panel of Figure 8. Unlike the bias in the total magnitude, however, the bias in using average corrected aperture colors is quite small. Over the range of sizes, $1 < r_h < 30$ pc, the $g-z$ aperture correction deviates by only $^{+0.004}_{-0.014}$ from the $r_h = 3$ pc fiducial. This shows that average corrected aperture colors can be considered essentially unbiased for the typical range of GC sizes, and they have the added advantage of smaller photometric errors for faint sources as compared to colors from model magnitudes.

When we further examine the effects of GC size on the $g-z$ aperture correction, we can see that at small sizes, the aperture correction to the color gets redder with increasing GC size due to the different sizes of the PSFs in F475W and F850LP. However, once GCs are larger than $r_h \sim 10$ pc, the color correction becomes bluer. This can be understood because at that point, we enter the regime where the size difference between the PSFs in the two filters is small compared to the size of the GC (10 pc $\sim 0''.125$ at Virgo). Therefore, GCs with increasingly larger sizes will not produce increasingly redder corrections. However, for studies where the objects of interest have sizes $r_h > 10$ pc, we recommend the use of model magnitudes and colors only.

Recently, there has been a claim that aperture magnitude biases as described above can masquerade as astronomical effects (Kundu et al. 2008). In particular, the claim is that trends in the mean colors of metal-poor GCs in the color–magnitude diagram can be explained by a correlation between GC size and luminosity. This cannot be true because our study of the color–magnitude relation of metal-poor GCs (Mieske et al. 2006a) uses model magnitudes and colors, which fully account for size-dependent aperture corrections. Moreover, as we have just shown, even if using average corrected magnitudes, the color biases for the average correction aperture colors would be much too small to account for the color–magnitude relation seen in GC systems.

4. COMPLETENESS

For many studies, it is important to know the level of completeness for a GC with a given set of properties. Unlike for point sources in blank fields where object magnitude is the primary parameter driving the detection probability, the ACSVCS data present more parameters than can affect completeness, namely the GC size (i.e., surface brightness), and the brightness of the background from the integrated light of the host galaxy. For a given magnitude, a GC with a larger size will be more difficult to detect, as will a GC projected onto the bright central regions of the host galaxy. We have run an extensive set of simulations to quantify the completeness of GCs as a function of their size (r_h), the background surface brightness (μ_b), and their total magnitude (m).

We created a suite of simulated GCs based on King models with $c = 1.5$ and sizes of $r_h = 1, 3, 6,$ and 10 pc. These simulated GCs were scaled to a random magnitude with $21 < z < 27$ mag and color $g-z = 1.1$. They were then inserted at semirandom positions in actual ACSVCS images for the galaxies VCC 1226 and 1833, the former being the brightest galaxy in the sample, and the latter being a dwarf galaxy with one of the lowest sky backgrounds. This allowed us to sample the full range of background surface brightnesses in the survey. Pixel fluxes for simulated GCs were given the appropriate random Poisson noise, and GC centering was randomly shifted by fractions of a pixel (as small as $1/60$ of a pixel). The positions at which they were placed in both the F475W and F850LP images were random, but they were limited to regions that were at least $3''.5$ from a real object, and at least $2''$ from the edge of the image. This was because our goal was to measure detection efficiency as a function of (m, r_h, μ_b) and not due to crowding or deblending, which does not present a problem for real objects.

The images with simulated GCs were then analyzed using the exact same pipeline as was used to produce the GC catalog. In total, we simulated 4,993,501 fake GCs across the full range magnitude, size, and background surface brightness. For each

Table 2
g-Band Completeness Curves

r_h	m_g	b_g									
		0.0518	0.1046	0.1491	0.2359	0.4023	0.7175	1.2535	2.1876	3.6955	6.0713
(")	AB mag										
0.0128	22.10	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.0128	22.11	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.0128	22.12	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.0128	22.13	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.0128	22.14	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.0128	22.15	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.0128	22.16	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.0128	22.17	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.0128	22.18	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.0128	22.19	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.0128	22.20	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.0128	22.21	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.0128	22.22	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.0128	22.23	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.0128	22.24	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.0128	22.25	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.0128	22.26	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.0128	22.27	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.0128	22.28	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.0128	22.29	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

(This table is available in its entirety in a machine-readable form in the online journal. A portion is shown here for guidance regarding its form and content.)

Table 3
z-Band Completeness Curves

r_h	m_z	b_z									
		0.0326	0.0917	0.1403	0.2354	0.4044	0.7457	1.3262	2.2897	4.3344	7.6519
(")	AB mag										
0.0128	21.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.0128	21.01	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.0128	21.02	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.0128	21.03	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.0128	21.04	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.0128	21.05	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.0128	21.06	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.0128	21.07	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.0128	21.08	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.0128	21.09	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.0128	21.10	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.0128	21.11	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.0128	21.12	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.0128	21.13	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.0128	21.14	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.0128	21.15	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.0128	21.16	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.0128	21.17	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.0128	21.18	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.0128	21.19	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

(This table is available in its entirety in a machine-readable form in the online journal. A portion is shown here for guidance regarding its form and content.)

of these GCs, we know whether or not they are detected, and we build completeness curves from this database. For each of the four sizes, and for each filter, we present the detection probability of a GC in 10 bins of background surface brightness roughly equally spaced in $\log(\text{flux})$, and in 0.1 mag steps in magnitude for $21 < z < 27$ mag. These completeness curves are presented in Table 2 for the g band and Table 3 for the z band. For each of the r_h values simulated (indicated in the first column) we tabulate the completeness curves as a function

of magnitude (indicated in the second column) and the 10 different background values (Columns 3–12, background values are specified in the header of these columns).

5. CATALOGS

5.1. Full Source Catalogs

In Table 4 we present our full catalog of sources satisfying the rough selection criteria presented in Section 2.6 of

Table 4
Photometric and Structural Catalog of Sources^a

VCC (1)	α (J2000) (2)	δ (J2000) (3)	$d_{\text{gal}}('')$ (4)	m_z (5)	$m_{z,\text{ap}}$ (6)	m_g (7)	$m_{g,\text{ap}}$ (8)	$r_{h,z}$ (9)	$r_{h,g}$ (10)	p_{GC} (11)	$E(B - V)$ (12)	b_z (13)	b_g (14)
1226	187.4439383	8.0000332	3.219	21.325 ± 0.050	21.271 ± 0.042	22.477 ± 0.057	22.477 ± 0.037	0.0301 ± 0.0041	0.0370 ± 0.0050	1.00	0.022	11.0150	7.5360
1226	187.4455059	8.0010107	3.437	22.700 ± 0.426	22.665 ± 0.115	23.797 ± 0.093	23.763 ± 0.084	0.0319 ± 0.0181	0.0283 ± 0.0084	1.00	0.022	10.4730	7.4200
1226	187.4436038	8.0011660	4.942	21.669 ± 0.043	21.608 ± 0.036	23.038 ± 0.046	22.988 ± 0.046	0.0229 ± 0.0042	0.0224 ± 0.0052	1.00	0.022	7.8560	5.5080
1226	187.4448279	7.9986757	6.244	20.916 ± 0.046	20.874 ± 0.028	22.297 ± 0.032	22.255 ± 0.028	0.0238 ± 0.0039	0.0286 ± 0.0027	1.00	0.022	6.3150	4.4670
1226	187.4469413	7.9997839	8.081	21.217 ± 0.031	21.164 ± 0.027	22.557 ± 0.029	22.511 ± 0.029	0.0225 ± 0.0033	0.0289 ± 0.0029	1.00	0.022	4.4850	3.2670
1226	187.4451763	8.0026245	8.123	22.770 ± 0.203	22.882 ± 0.097	23.611 ± 0.095	23.629 ± 0.061	0.0574 ± 0.0233	0.0411 ± 0.0088	1.00	0.022	4.6930	3.4040
1226	187.4454712	8.0026666	8.523	22.125 ± 0.087	22.114 ± 0.080	23.073 ± 0.038	23.032 ± 0.036	0.0348 ± 0.0086	0.0343 ± 0.0030	1.00	0.022	4.4050	3.2010
1226	187.4474911	8.0010406	9.987	22.309 ± 0.102	22.340 ± 0.050	23.501 ± 0.084	23.537 ± 0.054	0.0638 ± 0.0100	0.0419 ± 0.0075	1.00	0.022	3.2730	2.4030
1226	187.4423850	7.9987706	10.310	21.171 ± 0.036	21.147 ± 0.029	22.102 ± 0.020	22.076 ± 0.015	0.0325 ± 0.0035	0.0299 ± 0.0014	1.00	0.022	3.1990	2.3020
1226	187.4448252	8.0033496	10.594	21.498 ± 0.049	21.439 ± 0.037	22.893 ± 0.028	22.853 ± 0.025	0.0184 ± 0.0059	0.0148 ± 0.0026	0.94	0.022	3.5490	2.5860
1226	187.4464555	7.9979555	10.688	23.410 ± 0.578	23.437 ± 0.167	23.974 ± 0.512	23.926 ± 0.101	0.0077 ± 0.0898	0.0358 ± 0.0105	0.99	0.022	3.5530	2.5730
1226	187.4468670	7.9982008	10.933	19.889 ± 0.011	19.891 ± 0.009	21.288 ± 0.028	21.290 ± 0.021	0.0375 ± 0.0017	0.0349 ± 0.0018	1.00	0.022	3.4110	2.4890
1226	187.4440948	8.0035453	11.557	23.374 ± 0.101	23.318 ± 0.099	24.089 ± 0.123	24.095 ± 0.085	0.0326 ± 0.0079	0.0354 ± 0.0097	0.99	0.022	3.2620	2.3900
1226	187.4471733	8.0027498	12.050	22.594 ± 0.080	22.555 ± 0.066	23.934 ± 0.055	23.897 ± 0.052	0.0270 ± 0.0075	0.0141 ± 0.0056	1.00	0.022	2.6650	1.9680
1226	187.4433294	8.0034903	12.210	22.090 ± 0.045	22.024 ± 0.046	23.492 ± 0.052	23.445 ± 0.049	0.0219 ± 0.0052	0.0223 ± 0.0041	1.00	0.022	3.0790	2.2590
1226	187.4450661	7.9969838	12.365	22.292 ± 0.048	22.235 ± 0.045	23.785 ± 0.050	23.737 ± 0.053	0.0144 ± 0.0040	0.0270 ± 0.0058	1.00	0.022	2.9960	2.1770
1226	187.4479377	7.9985460	13.171	22.025 ± 0.027	21.956 ± 0.026	23.261 ± 0.109	23.220 ± 0.032	0.0249 ± 0.0035	0.0175 ± 0.0046	1.00	0.022	2.6190	1.9270
1226	187.4444844	8.0040618	13.196	21.760 ± 0.085	21.751 ± 0.036	23.246 ± 0.025	23.204 ± 0.024	0.0310 ± 0.0057	0.0260 ± 0.0033	1.00	0.022	2.7710	2.0380
1226	187.4416614	7.9980605	13.917	21.132 ± 0.022	21.109 ± 0.013	21.966 ± 0.044	21.938 ± 0.044	0.0227 ± 0.0020	0.0272 ± 0.0058	1.00	0.022	2.2020	1.5970

Notes.

Key to columns: (1) Galaxy VCC number; (2) and (3) J2000 right ascension (α) and Declination (δ) in decimal degrees; (4) Galactocentric distance in arcseconds; (5) z_{850} -band model magnitude obtained from the best-fits PSF convolved King model and an aperture correction as per Equation (9); (6) z_{850} -band average correction aperture magnitude inferred from a $0''.2$ aperture and an aperture correction as per Equation (10); (7) Same as (5) but for the g_{475} band; (8) Same as (6) but for the g_{475} band; (9) and (10) Best-fit half-light radii measured in the z_{850} and g_{475} bands, respectively; (11) Probability that the source is a GC according to the maximum likelihood estimate of our assumed mixture model (see Section 7); (12) Foreground $E(B - V)$ assumed for this source. The corrections for foreground reddening were taken to be $A_g = 3.634E(B - V)$ and $A_z = 1.485E(B - V)$ in the g and z bands, respectively (see Jordán et al. 2004); (13) Background in the z_{850} band (counts/s); (14) Background in the g_{475} band (counts/s).

^a Table 4 present the structural and photometrical catalog of all ACSVCS sources that satisfy the selection criteria presented in Section 2.6 in Jordán et al. 2004a (Paper II). To select a sample of bona fide GCs the sources should be restricted to those having $p_{\text{GC}} \geq 0.5$.

(This table is available in its entirety in a machine-readable form in the online journal. A portion is shown here for guidance regarding its form and content.)

Table 5
Catalog of Expected Contaminants^a

VCC (1)	$d_{\text{gal}}('')$ (2)	m_z (3)	$m_{z,\text{ap}}$ (4)	m_g (5)	$m_{g,\text{ap}}$ (6)	$r_{h,z}$ (7)	$r_{h,g}$ (8)	p_{GC} (9)
1226	8.127	21.563	21.797	22.413	22.396	0.0023	0.0192	0.65
1226	8.683	22.921	22.872	23.525	23.524	0.0150	0.0128	0.88
1226	9.106	20.545	20.483	21.252	21.258	0.0081	0.0145	0.54
1226	12.556	21.977	21.923	23.414	23.432	0.0124	0.0152	0.88
1226	15.859	23.186	23.116	24.067	24.066	0.0158	0.0128	0.89
1226	18.963	19.329	19.502	20.312	20.787	0.0051	0.0169	0.39
1226	18.966	22.586	22.818	23.941	24.055	0.0807	0.0525	0.94
1226	23.310	23.593	23.515	25.487	25.542	0.0154	0.0247	0.97
1226	23.385	23.312	23.550	24.597	24.811	0.0930	0.0775	0.44
1226	24.904	21.590	21.530	22.824	22.836	0.0119	0.0138	0.80
1226	25.750	19.101	19.182	20.782	21.209	0.0168	0.0226	1.00
1226	25.870	22.497	22.512	23.360	23.419	0.0331	0.0311	1.00
1226	26.637	22.120	22.197	23.049	23.497	0.0180	0.0230	1.00
1226	26.855	23.170	23.568	24.588	24.918	0.1098	0.0887	0.27
1226	27.742	20.256	20.205	20.757	20.760	0.0076	0.0159	0.56
1226	28.061	22.977	23.297	23.885	24.124	0.1301	0.0742	0.36
1226	28.708	24.031	23.970	24.610	24.576	0.0144	0.0206	0.92
1226	29.768	20.871	20.793	21.492	21.535	0.0134	0.0070	0.50
1226	30.516	22.317	22.247	23.521	23.576	0.0121	0.0269	1.00

Notes.

Key to columns: (1) Galaxy VCC number; (2) Galactocentric distance in arcseconds; (3) z_{850} -band model magnitude obtained from the best-fits PSF-convolved King model and an aperture correction as per Equation (9); (4) z_{850} -band average correction aperture magnitude inferred from a $0''.2$ aperture and an aperture correction as per Equation (10); (5) Same as (3) but for the g_{475} band; (6) Same as (4) but for the g_{475} band; (7) and (8) Best-fit half-light radii measured in the z_{850} and g_{475} bands, respectively; (9) Probability that the source is a GC according to the maximum likelihood estimate of our assumed mixture model (Section 7).

^a This table presents the expected contaminants in 17 control fields customized to each galaxy. It can be used to infer 17 times the expected contamination in any given GC sample.

(This table is available in its entirety in a machine-readable form in the online journal. A portion is shown here for guidance regarding its form and content.)

Paper II for all galaxies in the ACSVCS. The first column is the galaxy ID in the Virgo Cluster Catalogue (VCC; Binggeli et al. 1985; see Table 1 in Côté et al. 2004 for NGC and Messier equivalents). Columns 2 and 3 give the right ascension α (J2000) and declination δ (J2000) of each source and Column 4 gives the projected distance to the center of the host galaxy in arcseconds. Columns 5 and 6 give the total King model magnitude and the total magnitude inferred from a $0''.2$ aperture for the z_{850} band. These magnitudes have been dereddened as described in Section 2.7 in Paper II and have had aperture corrections applied as described in Section 3. Columns 7 and 8 give the corresponding quantities for the g_{475} band. Columns 9 and 10 give the best-fit half-light radii of the PSF-convolved King (1966) models in arcseconds for the z_{850} and g_{475} bands, respectively. The uncertainties do not include systematic uncertainties arising from the PSF modeling which can be estimated to be of order $\approx 0''.005$ (see Jordán et al. 2005). In order to convert the half-light radii to physical units, the SBF distances to our galaxies presented in Mei et al. (2007) can be used. Column 11 gives the value of p_{GC} for each source. Column 12 gives the adopted value of $E(B - V)$ which is taken from the DIRBE maps of Schlegel et al. (1998). Finally, Columns 13 and 14 give the galaxy plus “sky” background in counts/s present under each source in the g_{475} and z_{850} bands, respectively. These quantities are necessary to estimate the expected completeness using the data presented in Tables 2 and 3. We note that while formally we obtain the best-fit concentrations, we do not provide them here given that they are rather uncertain even for the most luminous GC candidates. In any case,

these quantities were used only for one galaxy (M87) in the work presented in Jordán et al. (2004) regarding the connection between low-mass X-ray binaries and GCs, and, as discussed in Sivakoff et al. (2007), the results of that work do not depend on the use of concentrations.¹⁶

We stress that to select a catalog of bona fide GC candidates from the full source list presented in Table 4, the sample needs to be restricted to sources satisfying $p_{\text{GC}} \geq 0.5$. We include sources with $p_{\text{GC}} < 0.5$ in Table 4 in order to allow the construction of GC samples using different criteria than those we have adopted in the ACSVCS. We caution though that parameters in Table 4 that are obtained by fitting PSF-convolved King (1966) models should be viewed only as rough indications for sources with $p_{\text{GC}} \lesssim 0.5$, as those sources are *not* expected to be well represented by King models in general given that they are most likely background galaxies.

5.2. Control Field Catalogs

In Table 5, we present our full catalog of contaminants satisfying the same selection criteria as the sources in Table 4 for all galaxies in the ACSVCS. These catalogs are obtained from 17 control fields that have been customized for each galaxy as described in Section 2.1.1. These catalogs therefore give 17 times the amount of contamination expected in the field of

¹⁶ Regarding this point, it is worth noting that Jordán et al. (2007c) have shown, using *HST*/ACS and *Chandra* observations of Centaurus A (=NGC 5128), that the King model concentrations are not a fundamental variable in determining the presence of low-mass X-ray binaries in GCs.

each target galaxy. These catalogs are useful in assessing the effects of the residual contamination for any study performed using the GC catalogs constructed from Table 4. We stress that background galaxies are not expected to be well-described by King models, and therefore the best-fit parameters presented in Table 4 are useful only for the purposes just mentioned. Additionally, note that the uncertainties in the photometric and structural parameters are calculated under the assumption that the objects can be described by King models and thus are not presented for objects in the contaminants catalogs where this assumption has no grounds.

The first column is the galaxy ID in the VCC (Binggeli et al. 1985; see Table 1 in Côté et al. 2004 for NGC and Messier equivalents). Columns 3 and 4 give the total King model magnitude and the total magnitude inferred from a $0''.2$ aperture for the z_{850} band. Columns 5 and 6 give the corresponding quantities for the g_{475} band. Columns 7 and 8 give the best-fit half-light radii of the PSF-convolved King (1966) models in arcseconds for the z_{850} and g_{475} bands, respectively. In order to convert the half-light radii to a linear distance, the SBF distances to our galaxies presented in Mei et al. (2007) can be used. Finally, Column 9 gives the value of p_{GC} for each source.

6. SUMMARY

We have presented the selection procedure for GC candidates in the ACS Virgo Cluster Survey, a survey of 100 galaxies in the Virgo cluster of galaxies. This procedure is based on model clustering methods which we briefly describe in the context of our survey.

We have additionally presented the determination of the aperture corrections for our GC candidates. Finally, we present the results of our photometric and structural parameter measurement for 20,375 objects which satisfy the rough selection criteria presented in Paper II in these series. This full source catalog contains 12,763 bona fide GC candidates which have a probability $p_{GC} > 0.5$ of being a GC according to our selection procedure. Additionally, we present catalogs of the contaminants expected to remain in such samples as deduced from observations of 17 control fields. These catalogs are presented as machine readable tables available for download from the electronic edition of the *Astrophysical Journal*. They are also available for download at the project's Web site <http://www1.cadc.hia.nrc.gc.ca/community/ACSVCS/>.

Support for program GO-9401 was provided through a grant from the Space Telescope Science Institute, which is operated by

the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS5-26555. E.W.P. acknowledges the support of the National Research Council of Canada's Plaskett Research Fellowship at the Herzberg Institute of Astrophysics. P.C. acknowledges additional support provided by NASA LTSA grant NAG5-11714.

Facilities: HST(ACS)

REFERENCES

- Bertin, E., & Arnouts, S. 1996, *A&AS*, **117**, 393
 Binggeli, B., Sandage, A., & Tammann, G. A. 1985, *AJ*, **90**, 1681
 Bruzual, G., & Charlot, S. 2003, *MNRAS*, **344**, 1000
 Chabrier, G. 2003, *PASP*, **115**, 763
 Côté, P., et al. 2004, *ApJS*, **153**, 223 (Paper I)
 Côté, P., et al. 2006, *ApJS*, **165**, 57 (Paper VIII)
 Côté, P., et al. 2007, *ApJ*, **671**, 1456
 Dempster, A. P., Laird, N. M., & Rubin, D. B. 1977, *J. R. Stat. Soc., Ser. B*, **39**, 1
 Ferrarese, L., et al. 2006a, *ApJS*, **164**, 334 (Paper VI)
 Ferrarese, L., et al. 2006b, *ApJ*, **644**, L21
 Ford, H. C., et al. 1998, in Proc. SPIE, 3356 Space Telescope and Instruments (Bellingham, WA: SPIE), 234
 Fraley, C. 1998, *SIAM J. Sci. Comput.*, **20**, 270
 Fraley, C., & Raftery, A. E. 2002, *J. Am. Stat. Assoc.*, **97**, 611
 Hasegan, M., et al. 2005, *ApJ*, **627**, 203 (Paper VII)
 Jordán, A., et al. 2004a, *ApJS*, **154**, 509 (Paper II)
 Jordán, A., et al. 2004b, *ApJ*, **613**, 279 (Paper III)
 Jordán, A., et al. 2005, *ApJ*, **634**, 1002 (Paper X)
 Jordán, A., et al. 2006, *ApJ*, **651**, L25
 Jordán, A., et al. 2007a, *ApJS*, **171**, 101 (Paper XII)
 Jordán, A., et al. 2007b, *ApJS*, **169**, 213
 Jordán, A., et al. 2007c, *ApJ*, **671**, L117
 King, I. R. 1966, *AJ*, **71**, 64
 Kissler-Patig, M., Jordán, A., & Bastian, N. 2006, *A&A*, **448**, 1031
 Kundu, A., Maccarone, T. J., & Zepf, S. E. 2008, in IAU Symp. 246, Dynamical Evolution of Dense Stellar Systems, ed. E. Vesperini, M. Giersz, & A. Sills (Cambridge: Cambridge Univ. Press), 408
 McLachlan, G. J., & Krishnan, T. 1997, *The EM Algorithm and Extensions* (New York: Wiley)
 Mei, S., et al. 2005a, *ApJS*, **156**, 113 (Paper IV)
 Mei, S., et al. 2005b, *ApJ*, **625**, 121 (Paper V)
 Mei, S., et al. 2007, *ApJ*, **655**, 144 (Paper XIII)
 Mieske, S., Hilker, M., Infante, L., & Jordán, A. 2006b, *AJ*, **131**, 2442
 Mieske, S., et al. 2006a, *ApJ*, **653**, 193 (Paper XIV)
 Peng, E. W., et al. 2006a, *ApJ*, **639**, 95 (Paper IX)
 Peng, E. W., et al. 2006b, *ApJ*, **639**, 838 (Paper XI)
 Peng, E. W., et al. 2008, *ApJ*, **681**, 197 (Paper XV)
 Salpeter, E. E. 1955, *ApJ*, **121**, 161
 Schlegel, D. J., Finkbeiner, D., & Davis, M. 1998, *ApJ*, **500**, 525
 Silverman, B. W. 1986, *Density Estimation for Statistics and Data Analysis* (London: Chapman and Hall)
 Sirianni, M., et al. 2005, *PASP*, **117**, 1049
 Sivakoff, G. R., et al. 2007, *ApJ*, **660**, 1246