Supplementary methods:

Image data collection

To collect our guenon images we used a Canon T2i fitted with Canon EF-S lenses of either 18-55mm, 55-250mm or 75-300mm depending on conditions. A circular polarizing filter was fitted to reduce the appearance of specular reflections off hair and glass. To photograph moving subjects we minimized exposure times by using a wide aperture (typically f/5.6) and high ISO rating (typically ISO 800). We avoided images containing objects that occluded a portion of the subject's face, those where the subject had a non-neutral facial expression, and photos from sequences where the subject remained in the same body position.

Additional image processing stages

Images were collected in RAW format and converted to linearized TIFFs using DCRAW [1]. To do this we required pixel values to be proportional to photon counts so we removed a slight remaining nonlinearity in sensor response to light intensity using a polynomial transform [2]. We then converted sensor RGB responses from the camera's colour space to guenon LMS colour space. Prior to collecting data we obtained estimates of the camera's spectral sensitivity empirically by imaging a reflectance standard through narrow band-pass filters and comparing camera responses to spectrometer readings. This, in conjunction with image exposure settings and estimates of guenon photoreceptor sensitivity [3,4], allowed us to undertake standard colour space transformation methods [2,5,6] to calculate the matrix transform between the camera's colour space and guenon colour space.

To correct for differences in ambient light and photographic exposure, and to model colour constancy processes in mammalian colour vision, we rescaled each channel with respect

to an estimate of each image's white-point. Because good practice is to care for captive primates 'hands-off' without any direct human contact, and because lighting was sometimes different inside and outside of enclosures, we were unable to always measure a reference standard under exactly the same illuminant as the subject. Instead we applied a computational colour constancy algorithm to images [7]. First, we cropped the subject out from areas of the background which either emitted light or were clipped above the dynamic range of the camera (typically the sky, artificial lights and specular reflections) and those which were under obviously different illumination from the subject (for example an artificially lit area outside a naturally lit enclosure). To correct for the colour of the illuminant we then made the grey world assumption [8] that the average colour of surfaces in a scene is grey, by equalizing the cone responses at each pixel based on the ratios of average long (L), medium (M) and short (S) wavelength photoreceptor responses. We then identified the most reflective surface in the scene (i.e. the maximum of all L, M and S values) and set this value as the white point, scaling other values accordingly. This perfect reflectance test assumes that there is a surface in the scene which reflects all light at a wavelength that one cone class is most sensitive to. We compared these white point estimates to measurements of the white patch on a Macbeth ColourChecker chart for images in which we considered there was good correspondence between the incident illumination of the subject and reflectance standard. We found a mean difference of 8.56% between the estimates and physical white points. This compares well with the mean 9.14% white point difference we have recorded between different images of the same physical standard taken in short succession from similar viewpoints under natural illumination [9].

To remove noisy pixels from the image caused by high ISO settings and standardize the texture of hair across images by normalizing for the effect of camera shake, we applied a circular

averaging filter of a diameter that gave a perceptually validated blur metric [10] score of 0.75, resulting in a set of equally slightly blurred images. The semi-automated process to separate the guenon's face from the background used spline interpolation to connect points selected by the user on the outline of the head. We selected the region bounded by the tips of hair originating along the jawline, around the outside of ears and over the top of the head.

Eigenface analysis

After transforming images to a standard reference frame for the eigenface analysis (main text), images were then further reduced in size to 98x75x3 pixels due to processing limitations. To conduct the eigenface analysis LMS responses were concatenated into a single vector for each image. The average face of the set (the guenon mean face for interspecific eigenfaces and a species mean face for intraspecific eigenfaces) was then subtracted from this and PCA was performed on each matrix of mean-shifted vectors to identify the dimensions on which face patterns vary between and within species. We examined scree plots to determine the appropriate number of eigenfaces to retain. Interspecific variation was well described by the first five eigenfaces. Intraspecific variation within each species was captured by between two and five eigenfaces, depending on the species and sample sizes. We then obtained inter and intraspecific eigenface scores for each original image by projecting them into inter and intraspecific eigenface space respectively.

Focal trait analysis

To extract focal traits from images, we applied the NSCT-SF-PCNN segmentation algorithm [11], which was implemented using the authors' MATLAB functions. In this model each image

pixel corresponds to a neuron in a network. Pixel brightness ((MW+LW)/2) is given as an external input to the network and neurons also receive input from neighbouring neurons. When the sum of inputs exceeds a threshold, the neuron gives a pulse output to neighbouring neurons, equivalent to an action potential in primate primary visual cortex. Outputs become synchronized based on the stimulus so that over a number of iterations the series of pulse outputs defines the regions in an image by forming autowaves (propagating waves that do not reflect or refract). The link arrangement parameter, which sets the number of neighbours each neuron is connected to, was set to six and the number of iterations to ten. Each nose-spot or eyebrow patch region was determined by selecting the largest segment in the region with only one sub-region from this map.

In order to describe the shape of focal traits, we undertook the elliptical Fourier analysis (EFA) method, which finds a set of harmonic coefficients for a function that defines shape as the sum of a series of ellipses. EFA was implemented in MATLAB using custom written functions. The outline of each nose-spot and eyebrow patch segment was recorded and converted to chain code [12]. Chains were initiated at the top left pixel and coding proceeded clockwise. As we were interested in both the separate and joint contribution of shape and size to classification performance, we normalized the size of each shape to the first harmonic. We examined scree plots and the accuracy of shape reconstructions to determine the number of components to include in further analyses. The number depended on the complexity of the shape and the extent of shape variation within a set and ranged from between two and five components.

Discriminant factor analysis:

To test whether there is potential information that could be used to categorise different categories using faces and focal traits we undertook discriminant factor analysis (DFA). Specifically, we used the pDFA approach [13], which uses random permutations of data to construct a null distribution of the probability of correct classification thus enabling non-independent data such as ours, where individual identity is an additional factor nested within species, sex or age identity, to be analysed appropriately. In addition to the procedure described in the main text, because DFA is sensitive to unbalanced data, the pDFA randomization process selects a number of samples equal to the number of cases in the smallest group. DFA also requires that the number of predictors is smaller than the number of cases in the smallest class of objects, though this was not an issue for any of our multipredictor models. We selected individual identity as a control factor when classifying age, sex and species and in individual identity classification models we would have included age or sex as a restriction factor [13] if we had found a significant age or sex classification model, however for our data no restriction factors were required, nor is a control factor appropriate. In total for both multipredictor models we ran three analyses for species identity (nose-spot, eyebrow patch and interspecific eigenface predictors), 19 for individual identity classification (4 species meeting inclusion criteria with nose spots + 4 species with eyebrow patches + 11 species' intraspecific eigenface scores), five pDFA analyses for sex classification (1 species meeting criteria with nose-spots + 1 species with eyebrow patches + 3 species' intraspecific eigenface scores), and four for age group classification (1 species meeting criteria with nose-spots + 1 species with eyebrow patches + 2 species' intraspecific eigenface scores).

For each pDFA analysis the null distribution was calculated by performing DFA on 1000 random data sets created following the permutation procedure [13]. For each permutation a

number of randomly selected cases in each group equal to the number of cases in the smallest group were selected for the cross-validation process. For species age and sex classification randomization occurs at two levels; randomization of each individual to the group, and a random selection of data from each individual to be used in the DFA. For individual identity classification randomization only occurs at the latter. Cross-validation was repeated 100 times, and the test statistic (percentage classification accuracy) results were averaged. The distribution of test statistics was then compared to the cross-validated classification accuracy (also repeated 100 times and averaged) of a DFA of the observed data.

Prior to running analyses we tested the assumptions of normality using Kolmogorov-Smirnoff tests and examination of trait score histograms and applied transformations as necessary. Homogeneity of variance and covariance matrices between response variable groups for each predictor were tested using Box's M test, all of which were significant at > 0.001, necessitating no further action [14]. We checked for multivariate normality and outliers by examining trait scores on a normal distribution and excluded those above and below the 97.5% and 2.5% quantiles.

References:

1. Coffin, D. 2008 DCRAW. **2**.

2. Stevens, M., Párraga, C. A., Cuthill, I. C., Partridge, J. C. & Troscianko, T. S. 2007 Using digital photography to study animal coloration. *Biol. J. Linn. Soc.* **90**, 211–237. (doi:10.1111/j.1095-8312.2007.00725.x)

3. Jacobs, G. H. & Deegan, J. F. 1999 Uniformity of colour vision in Old World monkeys. *Proc. R. Soc. B Biol. Sci.* **266**, 2023–8. (doi:10.1098/rspb.1999.0881)

4. Bowmaker, J. K., Astell, S., Hunt, D. M. & Mollon, J. 1991 Photosensitive and photostable pigments in the retinae of Old World monkeys. *J. Exp. Biol.* **156**, 1–19.

5. Westland, S. & Ripamonti, C. 2004 *Computational colour science using MATLAB*. Chichester: Wiley. [cited 2012 Aug. 10].

6. Stevens, M., Stoddard, M. C. & Higham, J. P. 2009 Studying primate color: towards visual system-dependent methods. *Int. J. Primatol.* **30**, 893–917. (doi:10.1007/s10764-009-9356-z)

7. Ebner, M. 2007 *Color Constancy*. John Wiley & Sons.

8. Buchsbaum, G. 1980 A spatial processor model for object colour perception. *J. Franklin Inst.* **310**, 1–26. (doi:10.1016/0016-0032(80)90058-7)

9. Dubuc, C., Allen, W. L., Maestripieri, D. & Higham, J. P. 2014 Is male rhesus macaque red color ornamentation attractive to females? *Behav. Ecol. Sociobiol.* (doi:10.1007/s00265-014-1732-9)

10. Crété-Roffet, F., Dolmiere, T., Ladret, P. & Nicolas, M. 2007 The Blur Effect: Perception and Estimation with a New No-Reference Perceptual Blur Metric. In *SPIE Electronic Imaging Symposium Conf Human Vision and Electronic Imaging*, pp. 64920I–64920I–11.(doi:10.1117/12.702790)

11. Qu, X., Yan, J., Xiao, H. & Zhu, Z. 2008 Image fusion algorithm based on spatial frequency-motivated pulse coupled neural networks in nonsubsampled contourlet transform domain. *Acta Autom. Sin.* **34**, 1508–1514. (doi:10.1016/S1874-1029(08)60174-3)

12. Freeman, H. 1961 On the Encoding of Arbitrary Geometric Configurations. *IEEE Trans. Electron. Comput.* **EC-10**, 260–268. (doi:10.1109/TEC.1961.5219197)

13. Mundry, R. & Sommer, C. 2007 Discriminant function analysis with nonindependent data: consequences and an alternative. *Anim. Behav.* **74**, 965–976. (doi:10.1016/j.anbehav.2006.12.028)

14. Tabachnick, B. G. & Fidell, L. S. 2001 *Using Multivariate Statistics*. 5th edn. Boston, MA: Allyn and Bacon.