Finding Spiral Structures in Images of Galaxies

B. D. Ripley and A. I. Sutherland

Email alerting service	Receive free email alerts when new articles cite this article - sign up in the box at the top right-hand corner of the article or click **here**

Finding spiral structures in images of galaxies

By B. D. Ripley and A. I. Sutherland

Department of Statistics, University of Strathclyde, 26 Richmond Street,
Glasgow G1 1XH, U.K.

Much recent work in statistical image analysis has been concerned with 'cleaning' images by a bayesian statistical analysis incorporating a prior model, which reflects the spatial structure of the image. In almost all cases this has involved a description of the image at pixel level. In this paper we take the process further, and develop a spatial stochastic process of objects present in the image. The general theory is given and applied to images of spiral galaxies, with the aims of producing better schematic reconstructions and of automatically classifying galaxies.

1. Bayesian image analysis

This has been the cornerstone of most recent statistical approaches to image reconstruction, 'cleaning' and de-blurring (Besag 1983, 1986; Geman & Geman 1984; Marroquin *et al.* 1987; Molina & Ripley 1989; Ripley 1986). A true image S is regarded as a member of finite-dimensional class of $\mathscr{S}$ of possible true images. The prior distribution $P(S)$ assigns preference amongst the class $\mathscr{S}$. An image Z is observed, almost always an array of intensities or colours at each rectangle of a $m \times n$ grid. (The rectangles are termed pixels.) The observation process is modelled by a probability distribution $P(Z|S)$ incorporating the noise, distortion, blurring and so forth of the observation process. Then inference about S is based on the posterior distribution $P(S|Z)$, which is given by

$$P(S|Z) \propto P(Z|S) P(S). \qquad (1)$$

The normalizing constant is not needed provided we are only interested in the relative weight given to 'true' images S by $P(S|Z)$.

In almost all examples (including all those in the references above), the description of $\mathscr{S}$ is again pixel based. That is, the true image is described by a $M \times N$ grid of either intensities or perhaps a grid of classifications. Most authors have taken the grids for Z and S as coincident, so that each pixel of a remotely sensed image is classified by landuse (Besag 1983, 1986), and a digital charge-coupled device (CCD) image of galaxies is reconstructed as a grid of intensities (Molina & Ripley 1989). From a traditional statistical perspective S is a realization of a two-dimensional stochastic process observed with noise and distortion, and the aim is to estimate the original realization S.

It has been commented frequently that the pixels of $\mathscr{S}$ are artificial, an artefact of the observation process. For example, in remote-sensing classification of landuse, a pixel could easily be made up of parts of two or more landuses. This problem can be ameliorated considerably by choosing smaller pixels, so that $M \times N$ is a much finer grid than $m \times n$ (an example is given in Ripley 1990). However, there remains the fundamental objection that the prior stochastic process is defined via an artefact; it
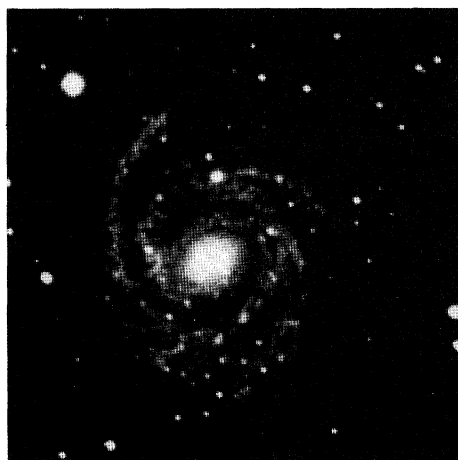
Figure 1. A 256×256 digital image with 256 grey levels. This is of galaxy NGC6814, classified *Sb*.
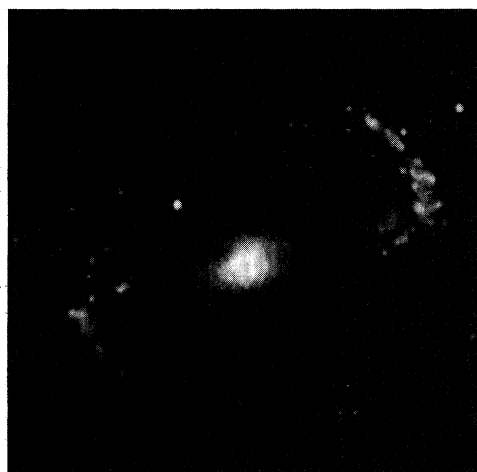

Figure 2. The *SBb(s)* galaxy NGC1300. This is the 'type' galaxy of that class. Image details as figure 1.

has proved difficult to relate the stochastic processes used at different grid sizes (Gidas 1989) let alone relate them to the underlying continuum. Nevertheless, for reasons of computational feasibility and convenience pixel-based priors are used, often rather successfully. Recent reviews are given by Besag (1989), Geman (1990) and Ripley (1988, 1990).

A different approach, that of describing objects in the image, has been pioneered by Grenander (1983) (cf. Chow *et al.* 1989). Our examples will be images containing stars and spiral galaxies. At our scale, stars are point sources and so can be described by location (on the celestial sphere) and apparent brightness. Spirals need a fuller description; two examples are shown in figures 1 and 2. Our description (elaborated in later sections) will be by the radius of the central disc, the length of the bar (if any)

and a description of a finite number of arms. Each arm is defined by a finite number of segments of fixed length together with the angles between segments. That describes a single spiral within a plane; an image then contains stars, and spirals transformed by a specified translation, scale, rotation in three dimensions followed by a projection onto the viewing plane. Although this amounts to a large number of parameters, there are in total tens or hundreds rather than the hundreds of thousands of (pixel) photon-counts which make up the raw data. Chow *et al.* (1989) worked with outlines of a human right hand. The outline is described by a series of arcs, about 36, forming a closed curve. Each arc is specified by a scaling and rotation; there is also a global transformation (translation and so on).

Such descriptions completely avoid including any part of the observation process. They result in a very appreciable data reduction, and provide descriptions that are very easy to interpret. They are also very suitable to use as the input to a classification scheme for galaxies. The end point of the data analysis might be the sort of catalogue a human interpreter might produce, a list of the form

at coordinate x, y there is a $SBab(r)$ galaxy of total brightness M.

Simpler descriptions underlying $\mathscr{S}$ also allow more information to be provided about $P(S\,|\,Z)$. When $\mathscr{S}$ is pixel-based, some extreme summary is needed. It has become conventional to show the mode of $P(S\,|\,Z)$ as the MAP (maximum *a posteriori*) estimator of S. Except in one very special case (Greig *et al.* 1989) the MAP estimator is too expensive to compute exactly, so in fact what is shown is the results of our best attempts at the implied optimization. One of the most-used methods of optimization is simulated annealing (Ripley 1987, §7.2; Aarts & Korst 1989). This depends on taking samples from

$$P_\lambda(S\,|\,Z) \propto [P(S\,|\,Z)]^\lambda$$

for large inverse temperature λ. Since λ is large, it is clear that P_λ concentrates on images S with high posterior probability, and hence images with maximal or near-maximal $P(S\,|\,Z)$ are produced more frequently. An alternative would be to show a few samples from $P(S\,|\,Z)$. With the methods of simulation used for pixel-based descriptions it is difficult to achieve sufficiently independent samples to give a representative view of $P(S\,|\,Z)$. With object-based descriptions samples from $P(S\,|\,Z)$ are easier to generate and tend to be much easier to interpret. It is also possible to estimate mean values for part of the description S, just by averaging over the samples, and so to show an estimated mean spiral galaxy, for example.

Classification

As noted above, galaxies are classified into about 30 types by their shapes (ellipticals, spirals, spirals with bars, tightness of arms and so on). This can also be incorporated into the bayesian framework. Let $T = 1, ..., t$ label the types. Then we need to specify $P(T = \tau), \tau = 1, ..., t$, the underlying frequencies of each type. The analysis described above for each type gives $P(S\,|\,T)$ and $P(Z\,|\,S)$. Then

$$P(S, T\,|\,Z) \propto P(Z\,|\,S)\,P(S\,|\,T)\,P(T) \qquad (2)$$

and, in principle, this can be summed over S to give $P(T\,|\,Z)$. Then a classical rule in statistical pattern recognition (Devijver & Kittler 1982) would be to classify the galaxy as the type maximizing $P(T = \tau\,|\,Z)$. In the terminology of statistical decision

theory this is the Bayes' rule for the zero-one loss function. One hopes that $P(T|Z)$ will be heavily concentrated on one value; otherwise it may be helpful to display the distribution over the more probable types.

One way to overcome the apparent computational difficulty in summing over S is to use simulation methods. Suppose we generate M (large) samples T_i from $P(T)$, then for each generate an S_i from $P(S|T)$. We then have M independent samples (T_i, S_i) from $P(T, S) = P(S|T)P(T)$. Accept each sample with probability proportional to $P(Z|S_i)$, ideally with probability

$$P(Z|S_i)/\max\{P(Z|S)\,|\,S \in \mathscr{S}\}. \tag{3}$$

Then we obtain $m \leqslant M$ samples from $P(S, T|Z)$, and the frequency of $T_i = \tau$ amongst these m estimates $P(T = \tau \mid Z)$. Since we do not need $P(T|Z)$ at all precisely, a small value of m will suffice. (This is fortunate, since the acceptance probabilities may well be small, and hence m very much less than M.) Unfortunately this method does not work with the iterative simulation methods we use in later sections. (They reject changes to samples sequentially with probability depending on ratios of $P(S|Z, T = \tau)$, and so the probability of rejection depends on T as well as S.)

2. Stochastic models for spiral galaxies

The prior stochastic processes used in pixel-based descriptions $\mathscr{S}$ are chosen to be Markov random fields (surveyed by Geman 1990), chosen principally for computational simplicity. Provided $P(Z|S)$ is defined locally, at each pixel s $P(S_s|S_t, t \neq s$ and Z) depends only on S and Z at nearby pixels. (A precise statement is given by Geman & Geman (1984, p. 728).) This enables iterative simulation methods for $P(S|Z)$ to be computed locally, that is using the information at a pixel and within a small neighbourhood of that pixel.

We have encountered only one approach to using Markov random fields for spirals, that of Seiden & Gerola (1982). This is very complex, and we prefer to use a more geometric description $\mathscr{S}$. Again, for computational reasons, we need to take the prior distribution on $\mathscr{S}$ as simple as possible. For example, Chow *et al.* (1989) describe S by a closed sequence of a fixed number of arcs. Each arc is derived from a template by a rotation and uniform scaling, giving a transformation g_i. The prior distribution is then a gaussian Markov chain on $\{g_i\}$.

Spiral galaxies show a wide range of variations (Sandage 1961; Goldsmith 1985; Sandage & Bedke 1988). Our description of a single galaxy has a number of components.

(*a*) A central *disc*, of random radius.

(*b*) *Bars* radiating from that disc, of random (possibly zero) length.

(*c*) *Arms* comprised of a finite number of line segments of fixed (pre-specified) length connected at their endpoints.

(*d*) A *point process* attaching bars or arms to the central disc, and arms to bars. This is a process of points on the circumference of the disc, with starting angles as marks.

(*e*) A *global transformation* corresponding to viewing the galaxy at an arbitrary location and random three-dimensional rotation, plus an overall rescaling over a limited range. (Galaxies normally are contained within a single image, and have apparent linear dimension at least a few percent of the size of the image.)

(*f*) Parts (*a*)–(*e*) define a *sketch* of the galaxy image. This is converted to a prior
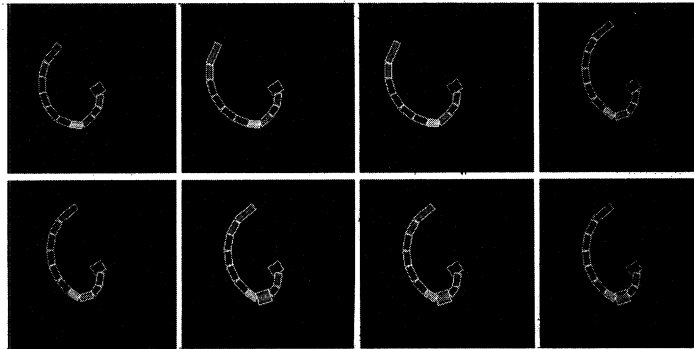
Figure 3. Eight simulations from the prior model for an arm, oriented in a similar way to one of the arms of figure 1. The white outlines indicate the width of the fitted grey-level surface. In these simulations the number of segments was kept fixed at 11.

on grey levels by assigning a cylindrical normal distribution to each line segment and a bivariate normal distribution within the central disc.

The distributions corresponding to all these components have to be completely specified, as do their dependencies. Wherever possible the components are taken as independent, for simplicity. For example, we might expect that the grey-level representations of adjacent line segments in an arm might be modelled by a dependent process, but this refinement was not introduced at first, and did not prove essential.

The description of an image consists of a list of objects, both spiral galaxies and stars. We can then regard the prior as a marked point process (Stoyan *et al.* 1987), where the marks are the (very complex) descriptions of the objects. For simplicity this marked point process is taken to be Poisson.

We have not yet considered all the levels of modelling an image in detail. The examples in later sections concentrate on fitting arms. An arm consists of a finite sequence of line segments of fixed length, joined at angles θ_i, and so is specified by $(n, \theta_1, \ldots, \theta_{n-1})$ where there are n segments and θ_i is the angle between segments i and $i+1$. This is a terminating Markov chain. Thus given the first r segments, $(r, \theta_1, \ldots, \theta_{r-1})$, there is a probability ϕ that there will be a segment $r+1$, and if there is, the angle θ_r between segments r and $r+1$ has a distribution depending only on θ_{r-1}. In our examples θ_i had a $N(\alpha\theta_{i-1}, \kappa_\theta)$ distribution, and $\phi \approx 0.9$ corresponding to an expected length of $1/(1-\phi) \approx 10$ segments. Some example are shown in figure 3, with n held constant at 11. The distribution of an arm depends heavily on θ_0, and this is specified as part of the 'attachment' point process.

3. Simulating samples from the posterior distribution

The simulation of a spiral from the prior is rapid since it is made up of a number of independent components (§2*a*–*f*). In principle, simulation of the posterior can be done by rejection sampling as at equation (3). To do so we must specify fully $P(Z|S)$. Studies of CCD images (Molina & Ripley 1989) showed that the noise was practically independent from pixel to pixel but had a variance which increased with intensity. The images of figures 1 and 2 were digitized from prints via a TV camera and are expressed as intensities on the scale 0–255.

For simplicity we worked with intensities on this scale 0–255, which certainly includes nonlinearities in the responses both of the film and the TV camera. We assumed independent additive gaussian noise variance κ at each pixel. From the description S we can obtain an 'ideal' image Y by averaging the grey levels of S over each pixel of Y. Then Z_{ij} is Y_{ij} plus additive noise. Thus

$$P(Z \mid S) \propto \exp\left[-\frac{1}{2\kappa}\Sigma\,(Z_{ij}-Y_{ij})^2\right], \tag{4}$$

where the sum is over all pixels. In our first experiments the grey-level height associated with an arm segment was taken to be uniform on $(0, 255)$. This produced many successful fits (acceptance of S) with arms which were almost black. In the same way that galaxies should not be too large or too small, 'arms' will not be regarded as such unless they are quite visible. Thus part (f) of the prior was redefined to have a linear distribution of intensity on $(0, 255)$. In this case the probability of acceptance is essentially zero unless the trial S has arms corresponding extremely well to those visible in the image, a very rare event by chance.

These difficulties may be overcome by iterative simulation (Ripley 1987, §4.7). Rather than simulate a whole galaxy at once and then accept (or more likely reject), we consider a change from S to S'. Then if

$$p = \frac{P(S' \mid Z)}{P(S \mid Z)} = \frac{P(Z \mid S)}{P(Z \mid S)}\frac{P(S')}{P(S)}, \tag{5}$$

the description S is replaced by S' with probability $\min(1, p)$. This process is repeated, building up a sequence of small changes to the description S. Under mild conditions this process converges, as the number of attempted changes increases, to give samples with distribution $P(S \mid Z)$. The conditions require symmetry of $S \rightarrow S'$ and $S' \rightarrow S$ and the ability to reach any point of $\mathscr{S}$ by a sequence of changes.

For simplicity, consider just one arm. A change could correspond to selecting one segment and altering its characteristics. If we alter the grey-level representation, the first term of (5) only depends on observations at pixels within that segment. If we alter the angle, all later segments on the arm are affected, but still the number of pixels affected, and hence the sum at (4), is a small part of the image. The net effect is that if the changes to S are chosen judiciously, p can be computed rapidly and will often be appreciable, so an acceptable proportion of proposed changes are made.

It may also be worth considering changes to more than one segment simultaneously. Changes to the angle of the first segment have a major effect on the overall fit and will rarely be accepted. Changes of the form

$$\theta_i \rightarrow \theta_i + \delta, \quad \theta_{i+1} \rightarrow \theta_{i+1} - \delta$$

shift segments $(i+2, ..., n)$ in a small parallel movement, and

$$\theta_i \rightarrow \theta_i + \delta, \quad \theta_{i+1} \rightarrow \theta_{i+1} - 2\delta, \quad \theta_{i+2} \rightarrow \theta_{i+2} + \delta,$$

leave segments $(i+3, ..., n)$ essentially unchanged.

The art of iterative simulation is to choose changes such that (5) is easy to calculate, their acceptance rate is high and iterations make rapid progress through the space $\mathscr{S}$ except where the fit to the observed image is very clear-cut. This is very much an art, at present depending heavily on watching experiments of iterative simulations.
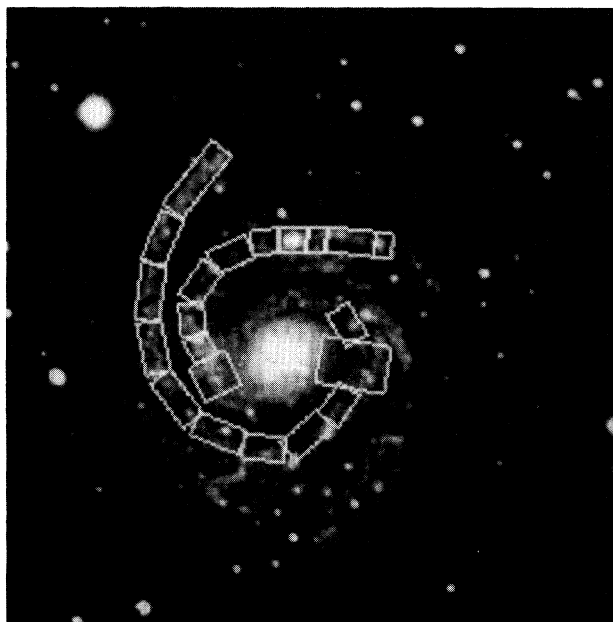
Figure 4. Fitted arms (samples from the posterior) for two of the arms of figure 1.
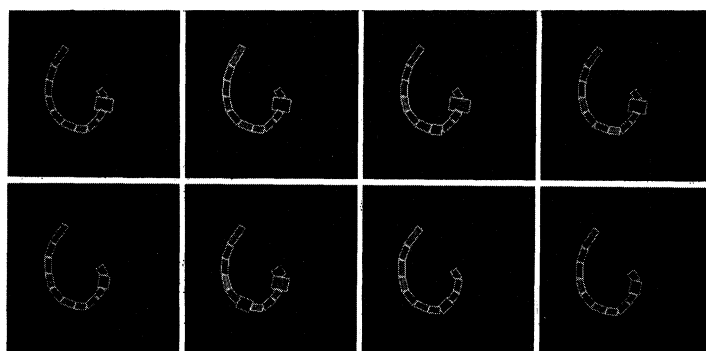

Figure 5. Eight posterior simulations for one of the arms of figure 1.

Examples

We illustrate just the fitting of arms to some of those in the galaxies NGC6814 and NGC1300 shown in figures 1 and 2. The initial position of the arm was taken from a manual sketch; the starting point and starting angle are held constant throughout the experiments.

Three of the four arms tried produced very satisfactory fits. It is clear from figures 4, 5 and 6 that the posterior distribution is heavily concentrated on one 'shape' for the arm, and there is serious uncertainty in the width of arm only where the intensity is low.

The lower arm of figure 2 caused difficulty. The posterior distribution is concentrated on an arm that bends round to overlap the upper arm. When we recall that arms are fitted independently this is not surprising. The lower arm shows a large
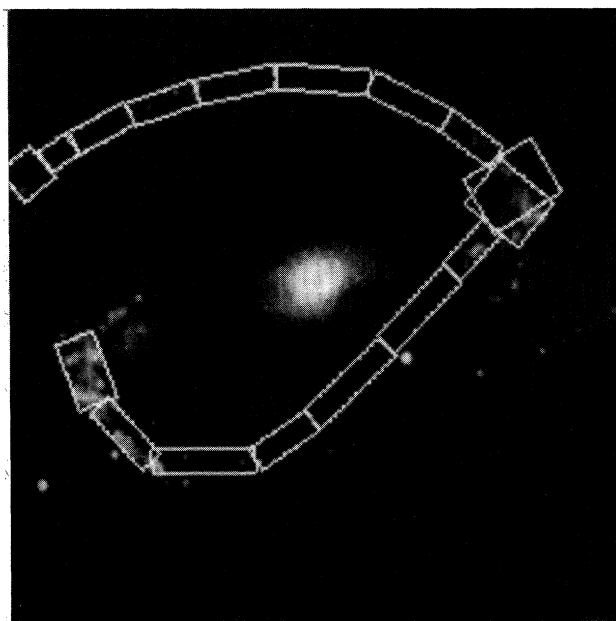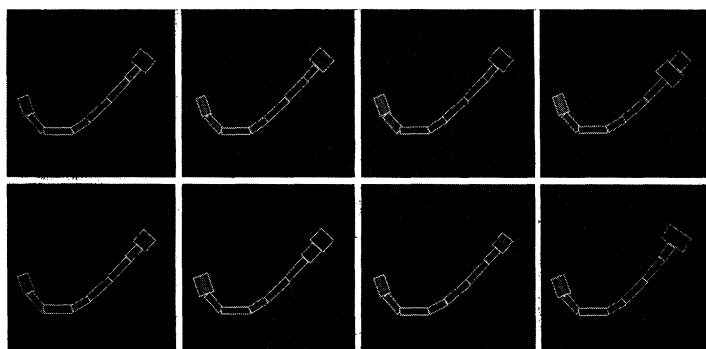
Figure 6. Fitted arms for figure 2.

Figure 7. A range of posterior simulations for the lower arm of figure 2.

'gap' in figure 2. Although the starting point for the iterative simulation was on what is to us the obvious arm, this has too low an intensity and is too diffuse, so rapidly the iterative process tries other positions of the arm before fixing on that shown. The posterior fits are shown with n fixed; with n allowed to vary a very short arm is chosen with an appreciable probability (figure 7).

These examples show some of the difficulties of identifying spirals. We tried a number of edge-detection techniques to extract the initial S automatically, completely without success. Yet the examples are fairly simple spirals; some galaxies have spiral arms that can be traced through many turns (Sandage 1961). We believe that it will only be possible to extract spiral structures by incorporating some global notion of shape, and that to regard the data as a sample from a spatial stochastic process of objects is perhaps the easiest way to do so.

References

Aarts, E. & Korst, J. 1989 *Simulated annealing and Boltzmann machines*. Chichester: Wiley.

Besag, J. 1983 Discussion of invited papers. *Bull. Int. Statist. Inst.* **50**, 422–425.

Besag, J. 1986 On the statistical analysis of dirty pictures (with discussion). *Jl R. stat. Soc.* B **48**, 259–302.

Besag, J. 1989 Towards Bayesian image analysis. *J. appl. Statist.* **16**, 395–407.

Chow, Y., Grenander, U. & Keenan, D. M. 1989 *HANDS. A pattern theoretic study of biological shapes*. Brown University: Division of Applied Mathematics.

Devijver, P. A. & Kittler, J. V. 1982 *Pattern recognition: a statistical approach*. Englewood Cliffs, New Jersey: Prentice-Hall.

Geman, D. 1990 Random fields and inverse problems in imaging. St Flour lectures. *Lect. Notes. Mathematics*. (In the press.)

Geman, D. & Geman, S. 1984 Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. *IEEE Trans. Pat. Analyt. Mach. Int.* **6**, 721–741.

Gidas, B. 1989 A renormalization group approach to image processing problems. *IEEE Trans. Pat. Analyt. Mach. Int.* **11**, 164–180.

Goldsmith, D. 1985 *The evolving universe*. Menlo Park, California: Benjamin/Cummings.

Greig, D. M., Porteous, B. & Seheult, A. 1989 Exact maximum *a posteriori* estimation for binary images. *Jl R. stat. Soc.* B **51**, 271–279.

Grenander, U. 1983 *Tutorial in pattern theory*. Brown University: Division of Applied Mathematics.

Marroquin, J. L., Mitter, S. & Poggio, J. 1987 Probabilistic solution of ill-posed problems in computational vision. *J. Am. stat. Ass.* **82**, 76–89.

Molina, R. & Ripley, B. D. 1989 Using spatial models as priors in astronomical image analysis. *J. appl. Statist.* **16**, 193–206.

Ripley, B. D. 1986 Statistics, images and pattern recognition. *Can. J. Statist.* **14**, 83–111.

Ripley, B. D. 1987 *Stochastic simulation*. New York: Wiley.

Ripley, B. D. 1988 *Statistical inference for spatial processes*. Cambridge University Press.

Ripley, B. D. 1990 The uses of spatial models as image priors. In *Spatial statistics and imaging* (ed. A. Possolo). *Inst. Math. Statist. Lecture Notes*. Hayward, California: IMS.

Sandage, A. 1961 *The Hubble atlas of galaxies*. Washington, D.C.: Carnegie Institution.

Sandage, A. & Bedke, J. 1988 *Atlas of galaxies useful for measuring the cosmological distance scale*. Washington, D.C.: NASA SP-496.

Seiden, P. E. & Gerola, H. 1982 Propagating star formation and the structure and evolution of galaxies. *Fundamentals Cosmic Phys.* **7**, 241–311.

Stoyan, D., Kendall, W. S. & Mecke, J. 1987 *Stochastic geometry and its applications*. Berlin, Chichester: Akademie & Wiley.

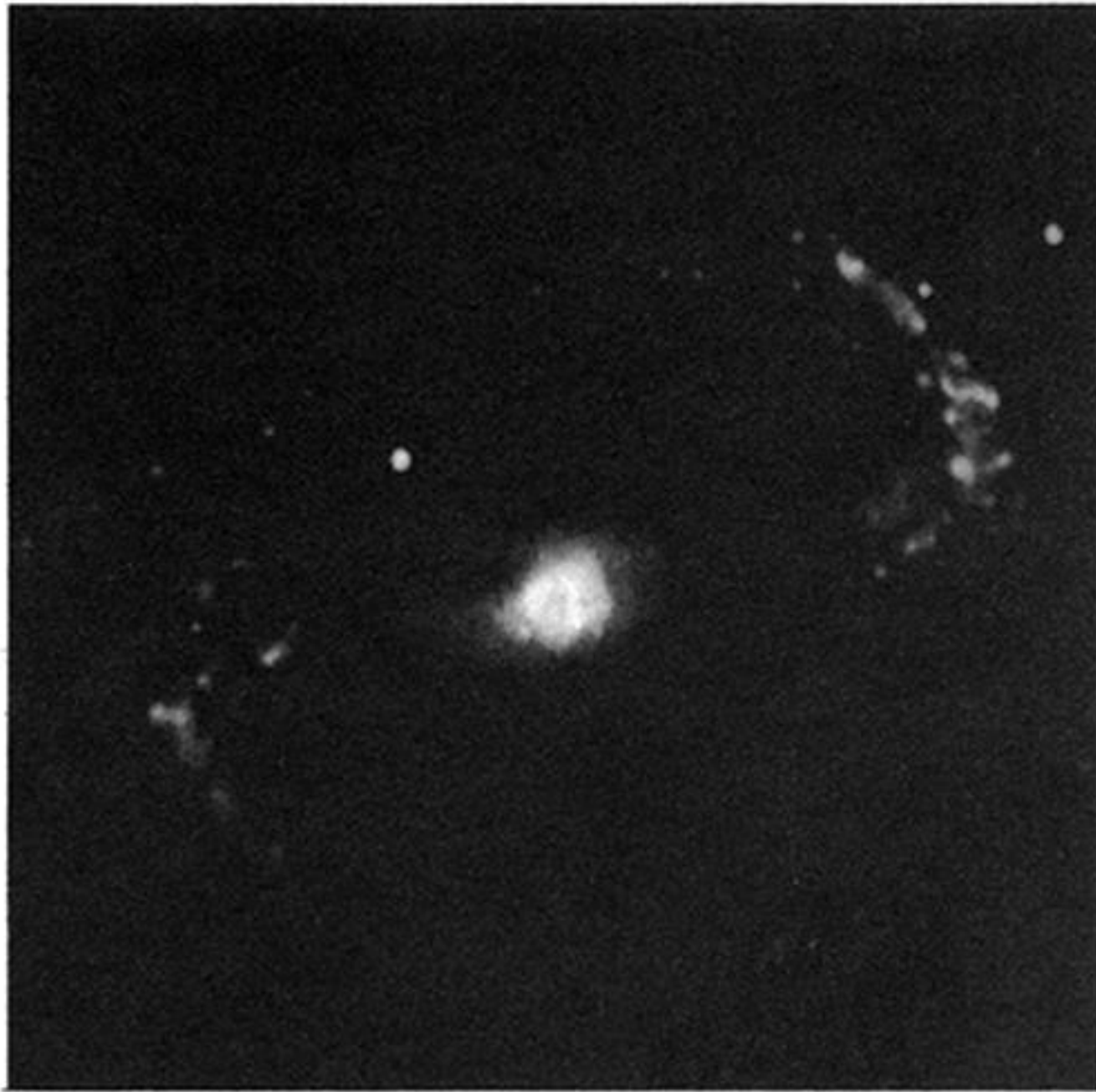gure 1. A 256×256 digital image with 256 grey levels. This is of galaxy NGC6814, classified Sb.

Figure 2. The *SBb(s)* galaxy NGC1300. This is the 'type' galaxy of that class. Image details as figure 1.
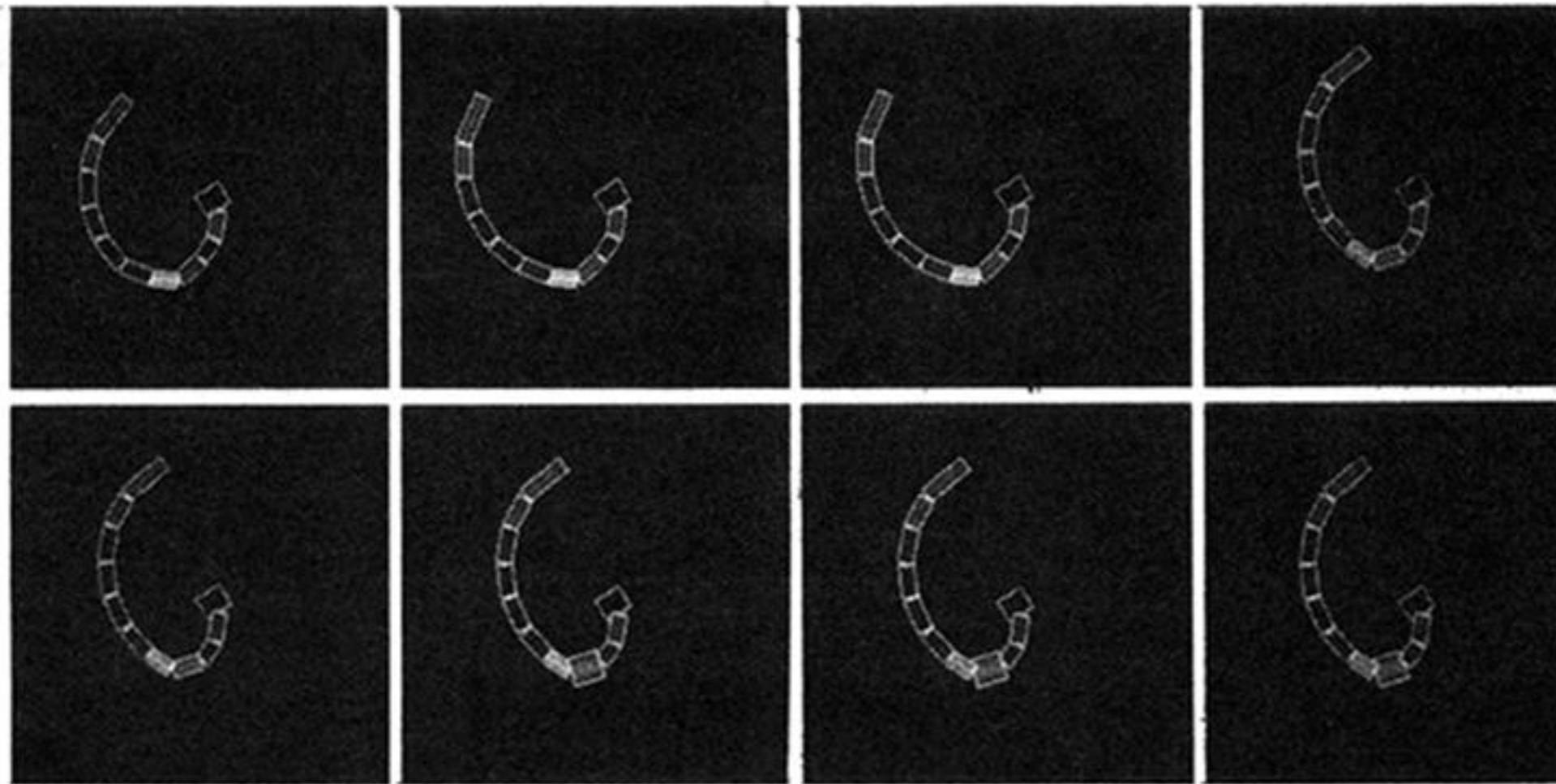
Figure 3. Eight simulations from the prior model for an arm, oriented in a similar way to one of the arms of figure 1. The white outlines indicate the width of the fitted grey-level surface. In these simulations the number of segments was kept fixed at 11.
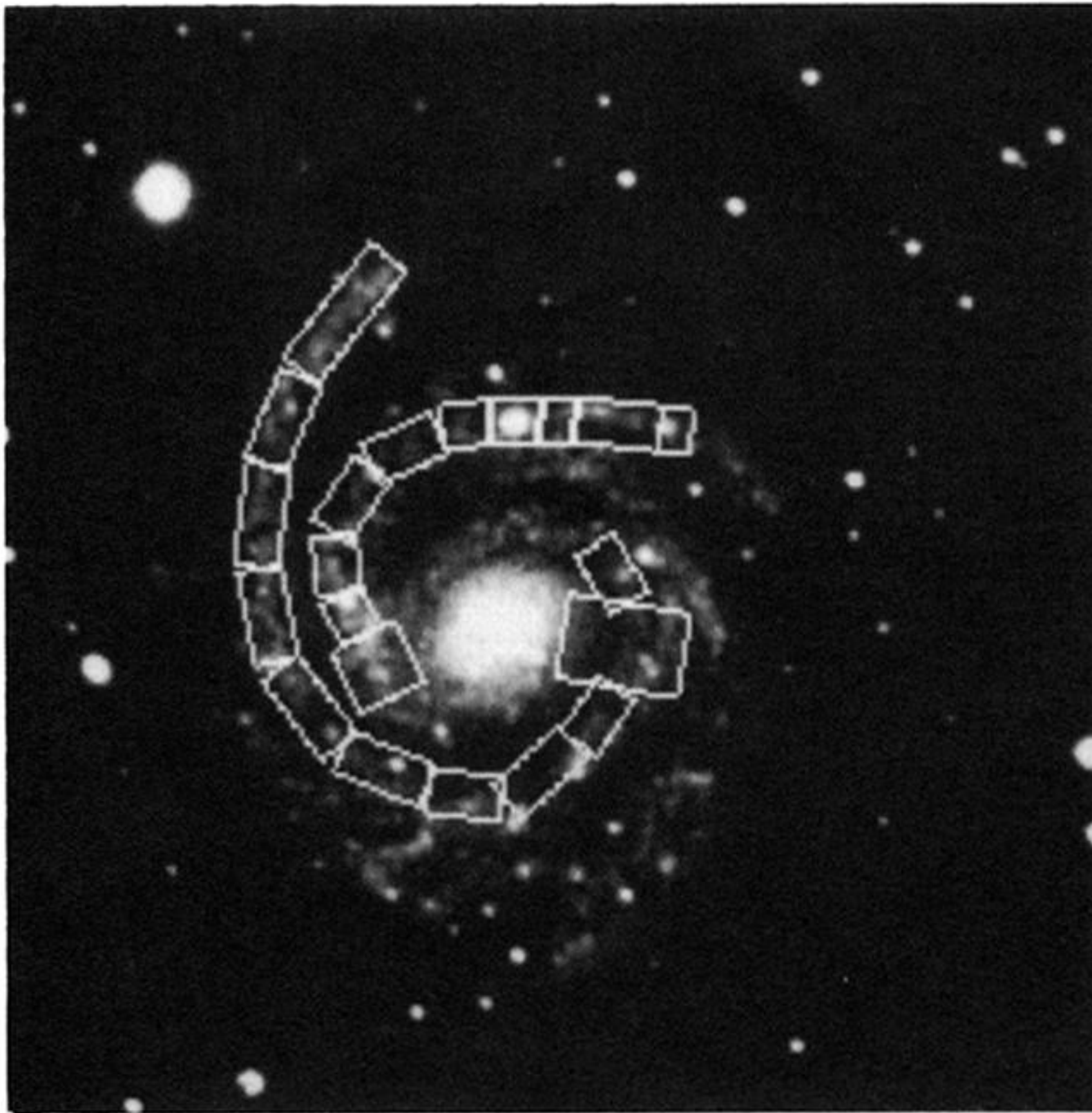
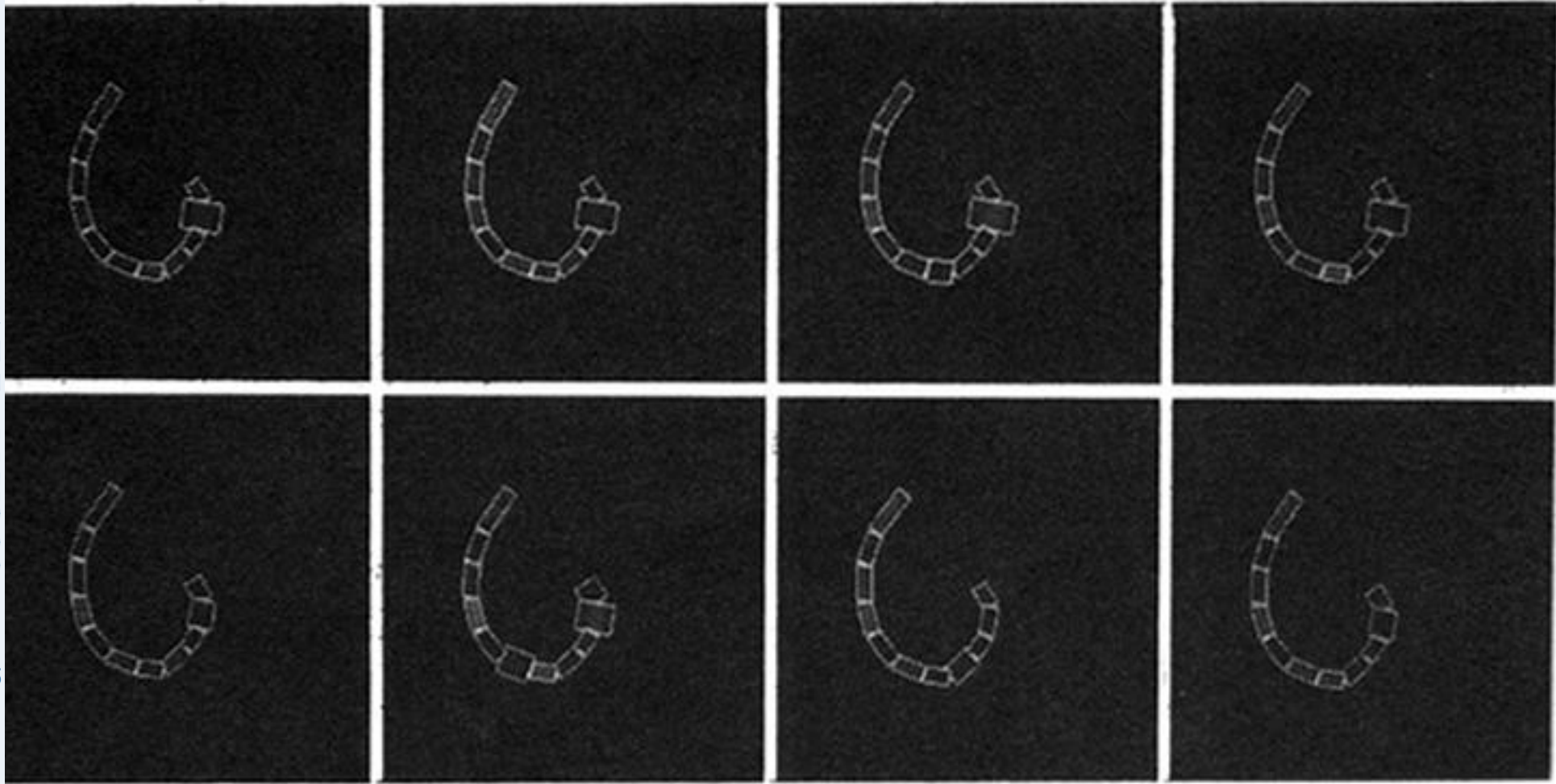Figure 4. Fitted arms (samples from the posterior) for two of the arms of figure 1.

Figure 5. Eight posterior simulations for one of the arms of figure 1.
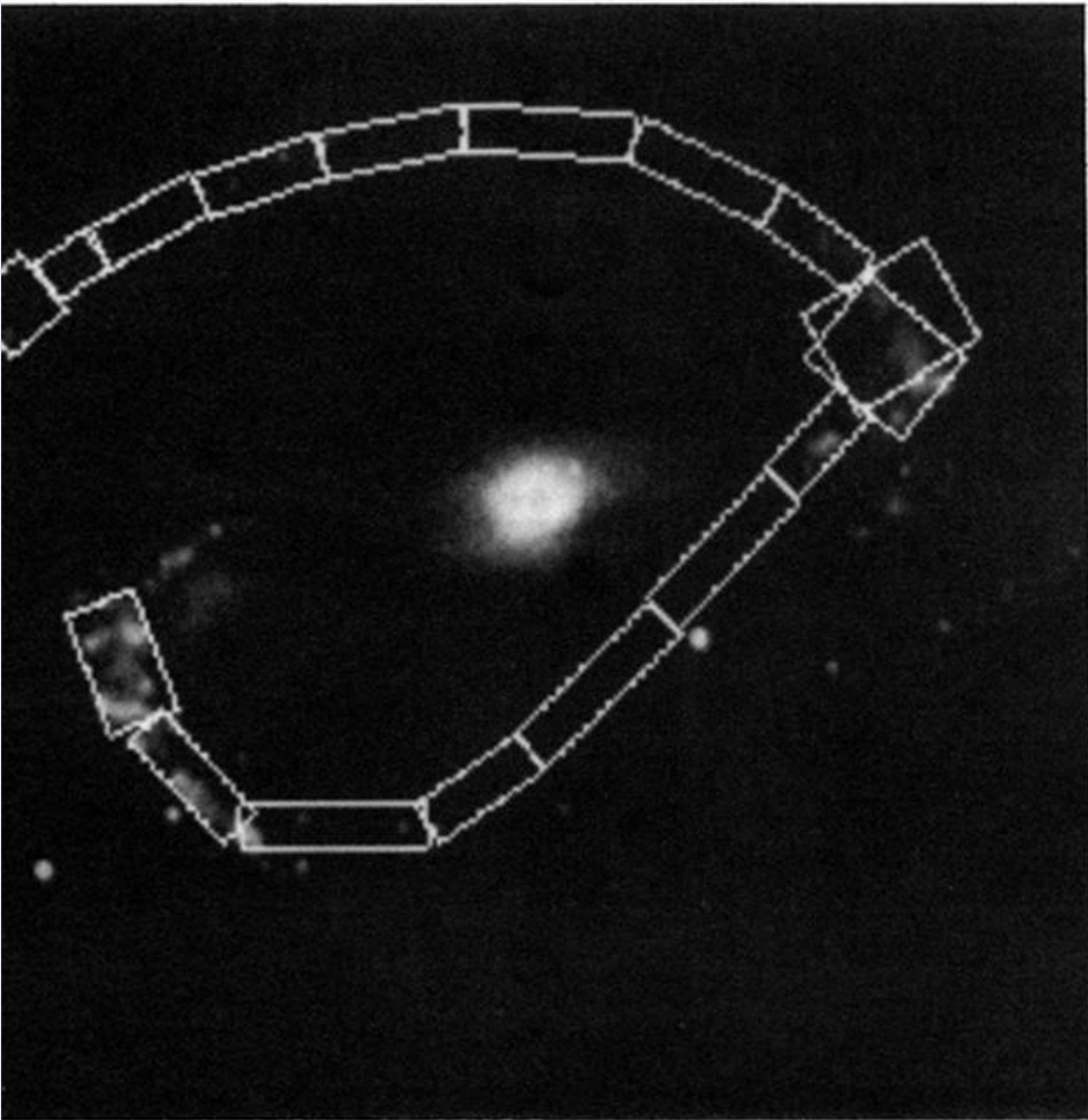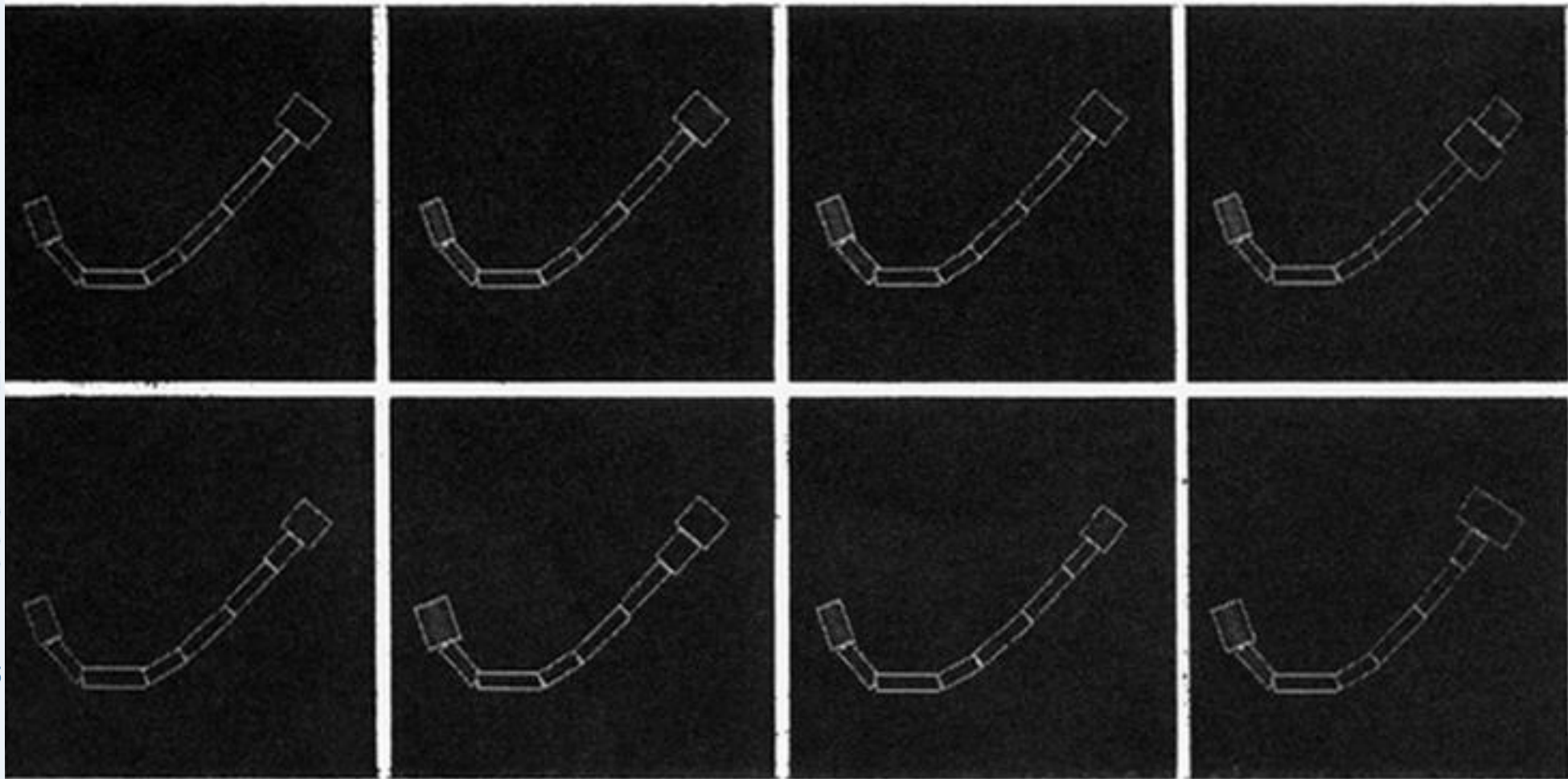
Figure 6. Fitted arms for figure 2.

Figure 7. A range of posterior simulations for the lower arm of figure 2.