

Bayesian Approach to Data Fusion

Introduction

- In Bayesian approach all forms of uncertainty are expressed in terms of probability.
- All unknowns are treated as random variables and that the knowledge of these quantities is summarized via a probability distribution.

Advantages of Bayesian Approach

- It is the only known coherent system for quantifying objective and subjective uncertainties.
- It provides principled methods for the model estimation and comparison and the classification of new observations.
- It provides a natural way to combine different sensor observations.
- It provides principle methods for dealing with missing information.

Bayesian Statistics

- Classical statistics:
 - Probability is interpreted as a properly defined limit of the frequency of events when an experiment is repeated infinitely.
- Bayesian statistics:
 - Probability is treated as a *Degree of Belief* (DoB).
 - Probability is a measure for uncertainty, it is a probabilistic measure to describe what is known.

A Simple Example: Completion of a Doctoral Thesis

A PhD student gives his supervisor a *precise time*, T , for the completion of his thesis.

- **Classical statistics:** It is highly likely for the student to finish his thesis at $t=T$.
- **Bayesian statistics:** Based on Bayes' rule , and prior information the estimate $t \approx T$ is *highly unlikely*
 - Priori information: a student working under pressure will almost certainly underestimate the time required to complete his/her thesis.
 - The likelihood function is only high for $t=T$, but *a priori probability* that T is a reasonable date, given the student is working under pressure, is low.

Bayesian Approach for Data Fusion

- Three stages of Bayesian approach:
 - **Probability Model:** creating a joint probability distribution that captures the relationship among all the variables under consideration.
 - **A Posterior Distribution:** summarizing all information regarding the different quantities of interest in a set of *a posteriori* distributions.
 - **Model Selection:** evaluating the appropriateness of using a given model.

Formulating Data Fusion Problem Based on Bayesian Approach

- Observation: M_i , $i=1,\dots,n$ of the same unknown Object Z
- Goal: estimate data $\hat{Z}$ representing the fusion of all M_i s

Likelihood function: $P(M_i | Z)$

Prior knowledge about Z : $P(Z)$

- If the data are observed independently, then we have

$$P(M_1, \dots, M_n | Z) = P(Z) \dots P(M_n | Z)$$

- According to the Bayes' rule, the posterior probability is

$$P(Z | M_1, \dots, M_n) = P(M_1, \dots, M_n | Z) P(Z) / P(M_1, \dots, M_n)$$

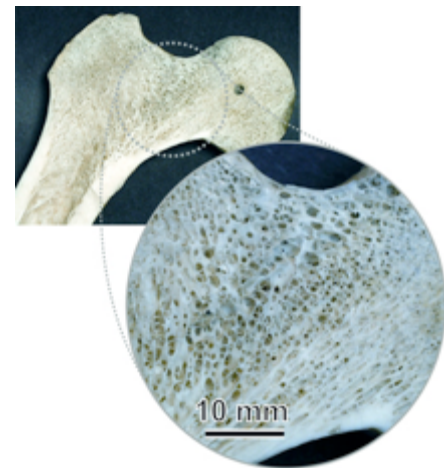
Denominator is the normalization factor.

Scientific Image Fusion for

Porous Media Image Reconstruction

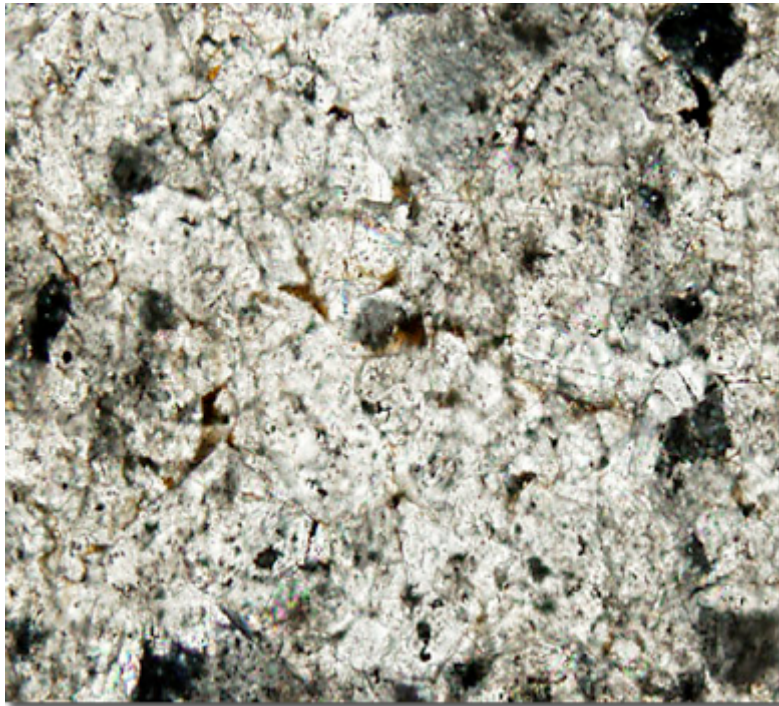
Porous Media

- Porous media:
 - The Science of water-porous materials such as cement, concrete, cartilage, bone, wood, and soil.
 - A solid material containing pores.
- Pore: a void inside a solid structure or solid material.

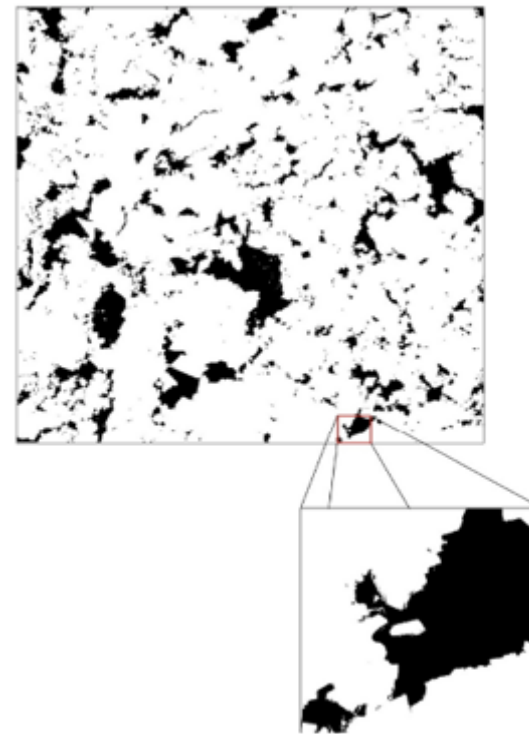


Examples of Porous Media

Carbonate Rock

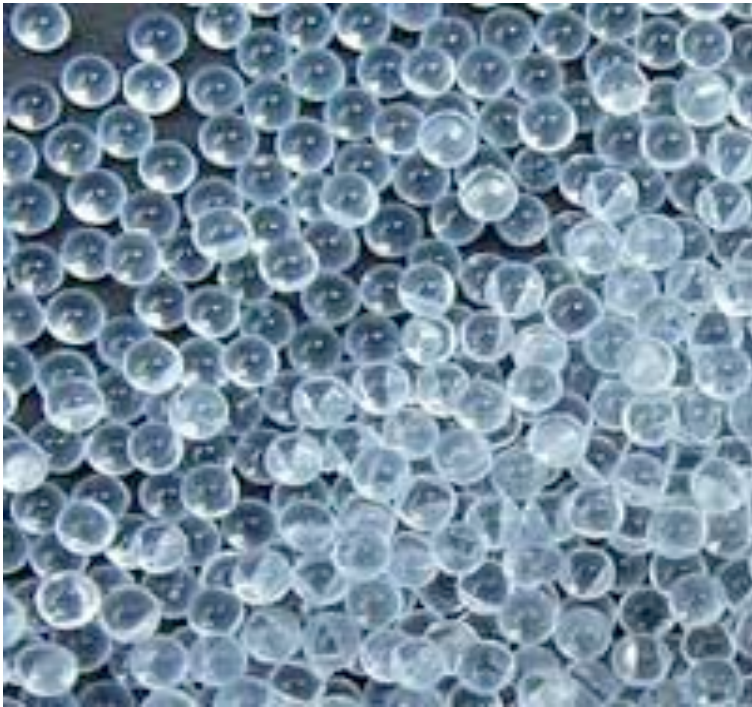


Two-dimensional Microscopic view
Pore is shown as black and Solid as white

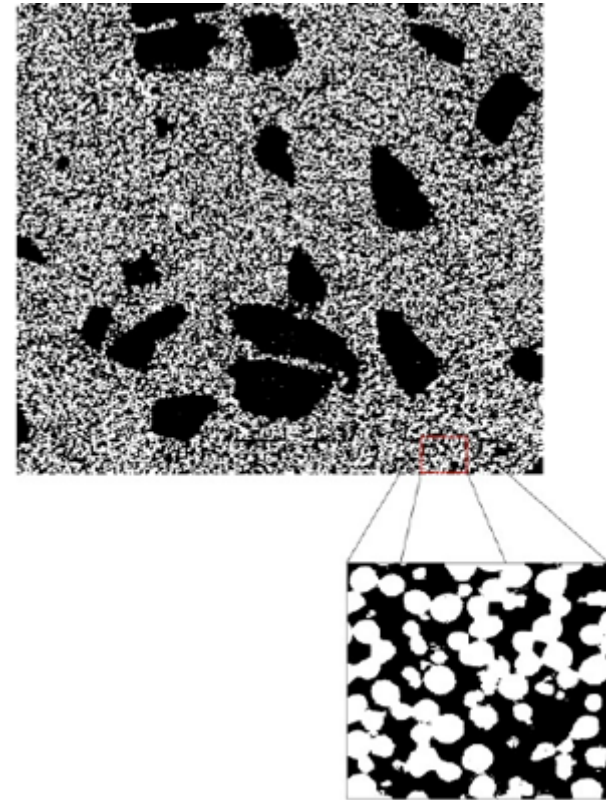


Examples of Porous Media

Sintered Glass Spheres

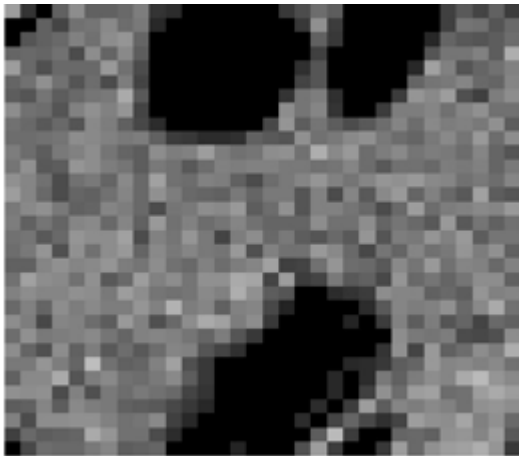


Two-dimensional (2D) Microscopic view
Pore is shown as black and Solid as white



Two dimensional Low Resolution Images of Porous Media

Single slice of three-dimensional (3D) low resolution measurement (observations) obtained by MR imaging



Sintered Glass Spheres



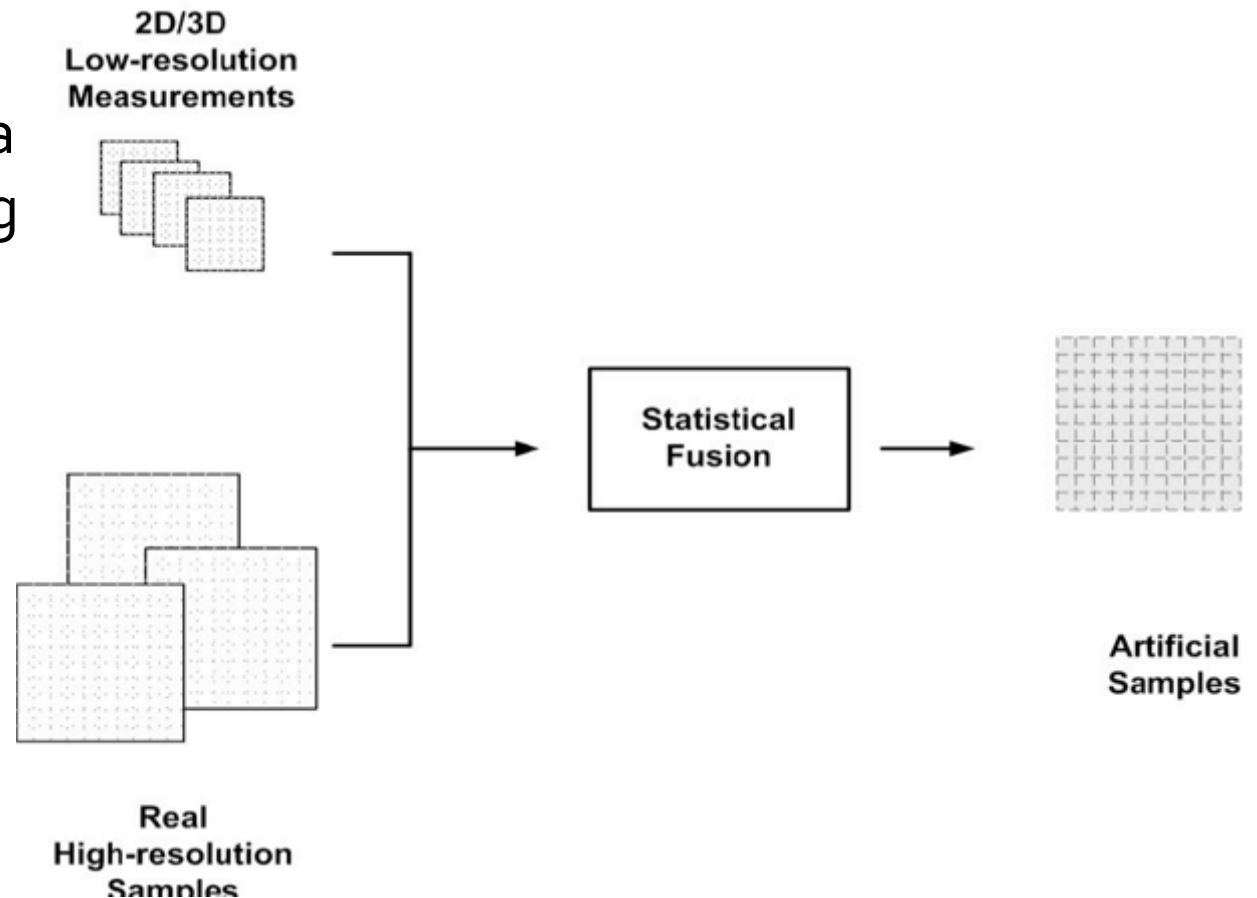
Carbonate Rock

Motivation for Data Fusion: Porous Media Analysis

- Many high-resolution 2D and 3D samples are required to study porous media properties such as pore structures, permeability, porosity and fluid transfer.
- Real measured 2D and 3D samples are problematic
 - 2D high resolution images of porous media requires cutting, polishing, and exposure to air
 - 3D samples generated by MR imaging only resolve low-resolution i.e. large-scale pore structures
- Artificial 2D and 3D image samples of porous media are generated using a *porous media reconstruction process*.

Porous Media Reconstruction

- Defining or Learning a statistical model using real 2D and 3D samples.
- Generating artificial samples from the statistical model.

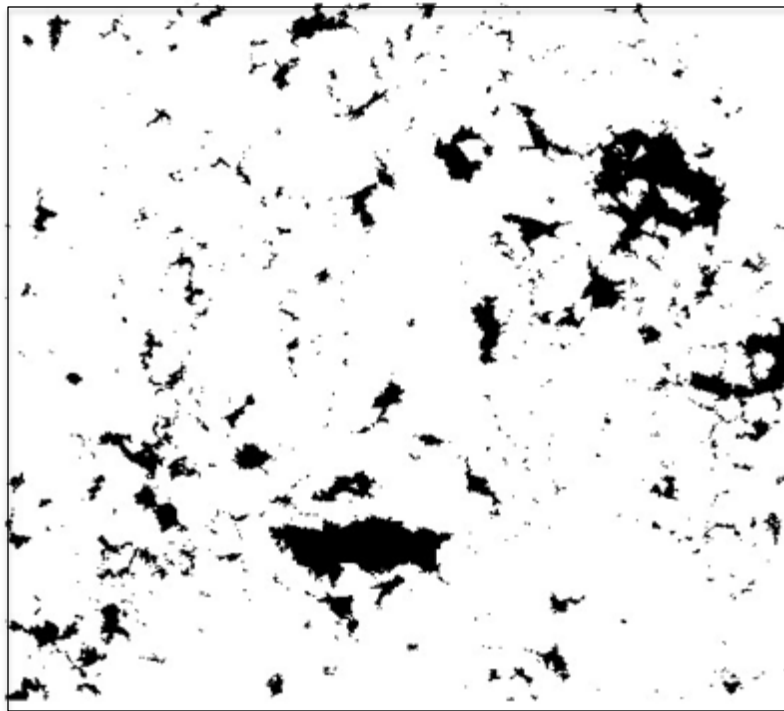


Porous Media Reconstruction

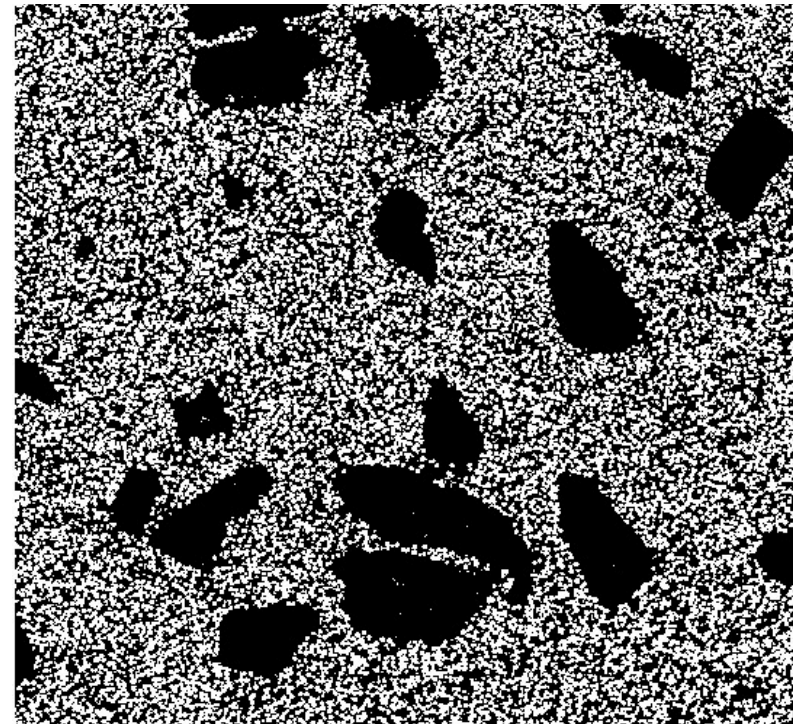
- Two approaches to generate artificial samples
 - **prior synthesis**: purely based on statistical prior models
 - **posterior synthesis**: based on coupling the prior model with measurements

Measured Data of Porous Media

2D high resolution images captured from optical microscopic imaging with $8\mu\text{m}$ resolution (binarized)



Carbonate Rock



Sintered Glass Beads

Measured Data of Porous Media

Slices of 3D Low resolution images
Showing the Large pores and an average of small
pores (Gray Scale Images)



Carbonate Rock

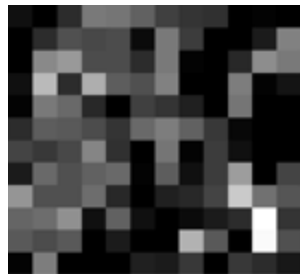


Sintered Glass Beads

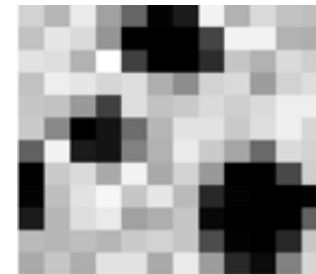
- The images are 30 times smaller in resolution than the high-resolution images.
- The measurements can be obtained by MR imaging.

Measured Data of Porous Media

Slices of 3D Low resolution MR images
Showing the surface-to-volume ratio, proportional
to pore size distribution (Gray Scale Images)



Carbonate Rock



Sintered Glass Beads

- The images are 30 times smaller in resolution than the high-resolution images.
- The measurements can be obtained by MR imaging.

Challenge

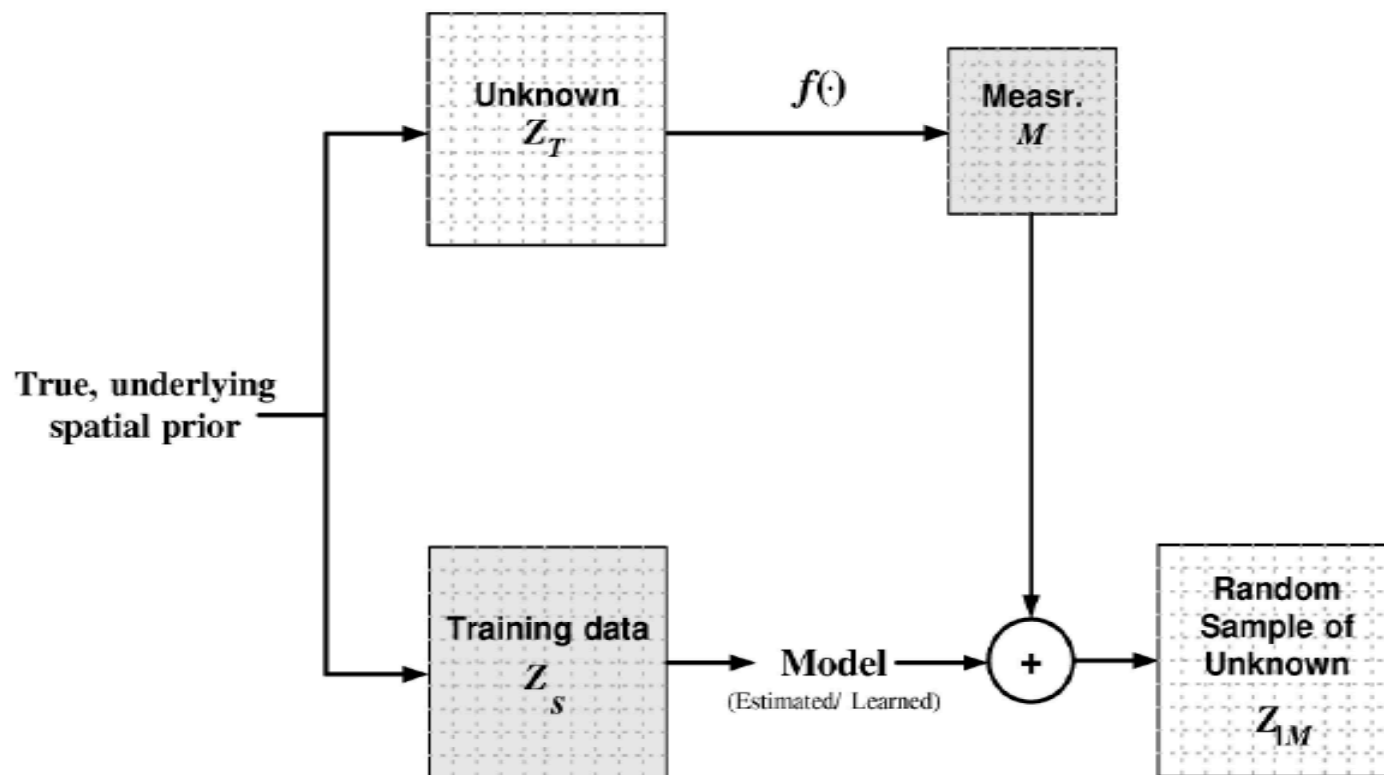
- **Goal:** fusing high resolution samples with low resolution measurements
- Why existing image fusion techniques are not practical?
 - many of such techniques solve an estimation problem instead of sampling (generating artificial samples)
 - we do not have multiple porous media samples from same scene so we can not use image super-resolution methods
 - most methods are for continuous-value images rather than discrete ones (binary images)
- We use a **Bayesian approach** to solve the fusion problem.

Applying Bayesian Approach

- Key point in Bayesian approach:
 - our prior knowledge (belief) regarding the true phenomenon under study can be modeled as a probability distribution.
- Evaluating the probability of a hypothesis in a Bayesian approach:
 - specifying some prior probability $P(\cdot)$, based on a priori knowledge or belief about the phenomenon (Z) under study,
 - updating the belief in the light of new and relevant observations (M) of the true phenomenon.

Bayesian Approach for the Fusion Problem

- The prior probability $P(Z)$ is combined with observations (measurements), M , to form a posterior probability, $P(Z|M)$.



Problem Definition

- For the statistical fusion problem, we are left with:
 - Defining the **prior model**, reflecting the statistical features and characteristics of the phenomenon under study
 - Defining the **observation (forward) model**, i.e. how the observations (measurements) are generated (the sensor model)
 - Defining a model (**posterior model**) to fuse the prior model with the observations;
 - **Generating samples** from the posterior model (probability) to obtain artificial image samples of porous media

Statistical Posterior Model

- The proposed model is based on the Gibbs probability distribution:

$$P(Z) = \frac{e^{-H(Z)/T}}{\Phi}$$

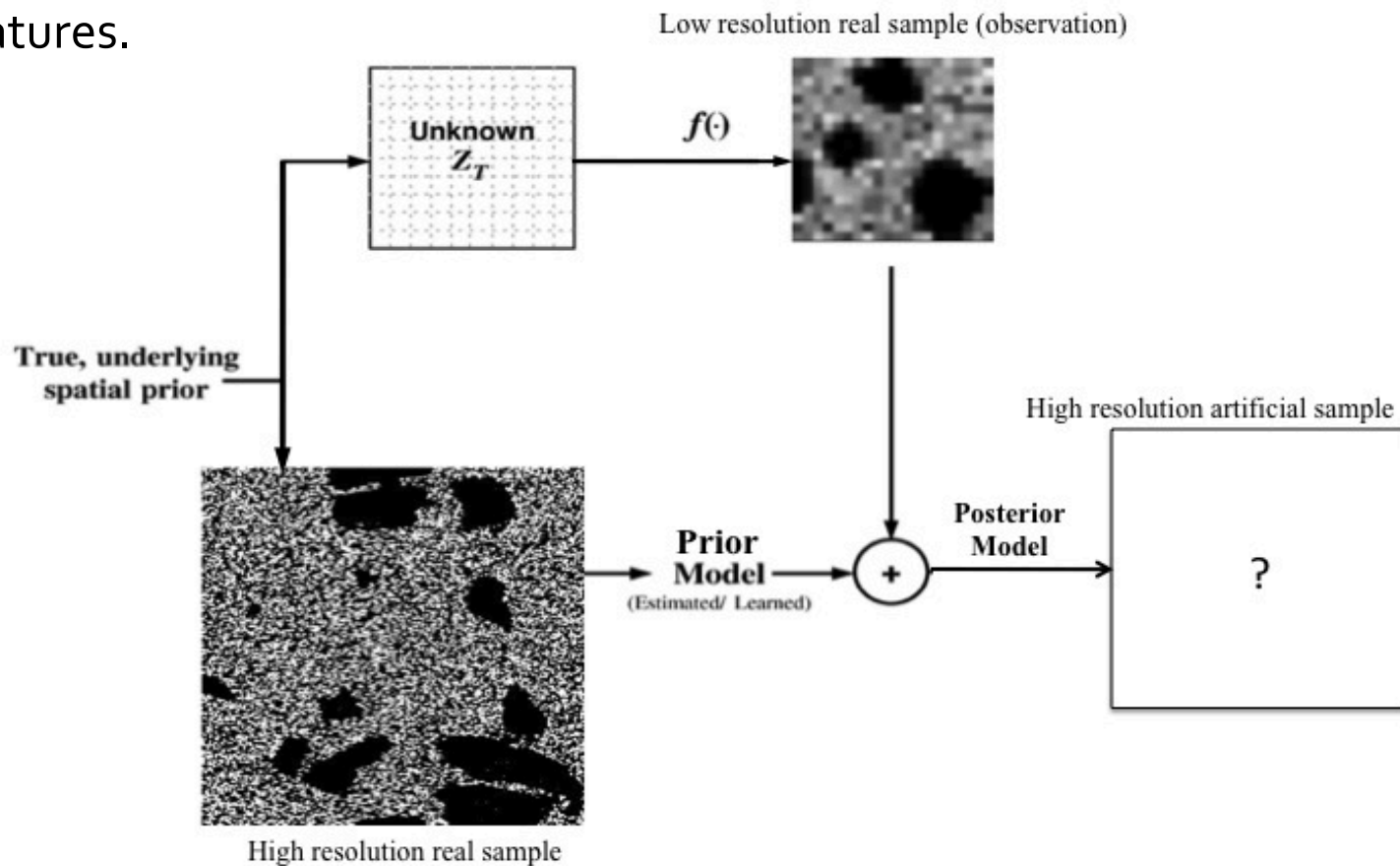
- $H(Z)$ is the prior energy function and Φ is a normalization factor.
- Given observations M , the posterior probability distribution is:

$$P(Z | M) = \frac{e^{-H(Z|M)/T}}{\Phi_M}, \quad H(Z | M) = H(Z) + \alpha J(Z; M)$$

- J is the observation (forward) model, describing how the observations are measured and generated.
- α balances the contributions of prior and observations, and J is the constraint.

Posterior Model

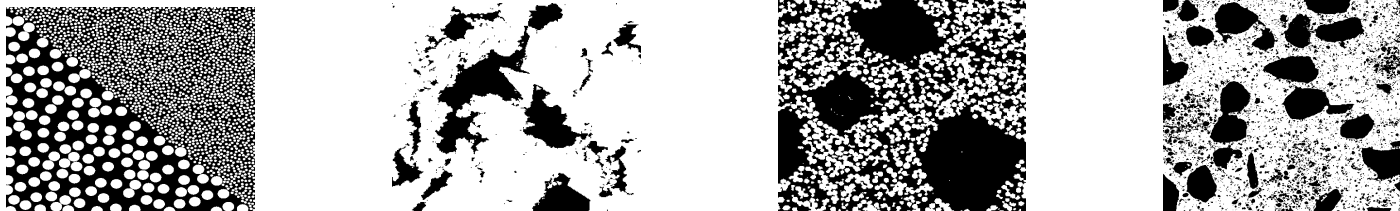
- The real low resolution and high resolution samples do not come from the same scene but from the same material with the same statistical features.



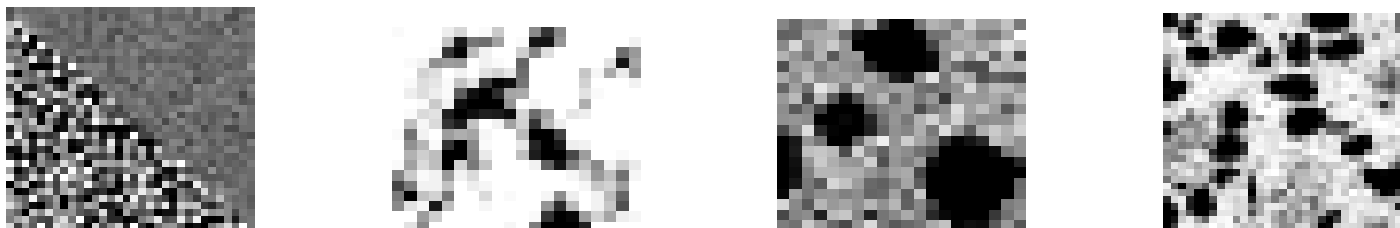
Posterior Sampling

- We are interested in generating random samples of porous media.
- Having both the advantage of prior model and measurement, we propose to do Posterior Sampling.
- Computing the normalization factor in $P(Z | M) = \frac{e^{-H(Z|M)/T}}{\Phi_M}$, i.e. Φ_M is almost impossible in practice.
- We use Markov Chain Monte Carlo (MCMC) methods along with simulated annealing to generate samples indirectly from the posterior energy.
- The posterior sample is similar to estimates in densely-measured area, and a random synthesis in those area not constrained by measured values.

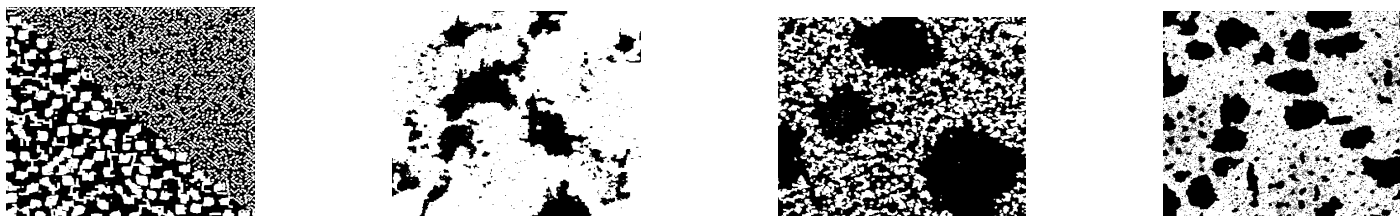
Results



(a) Original data

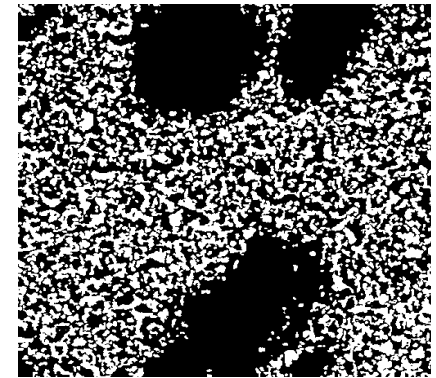
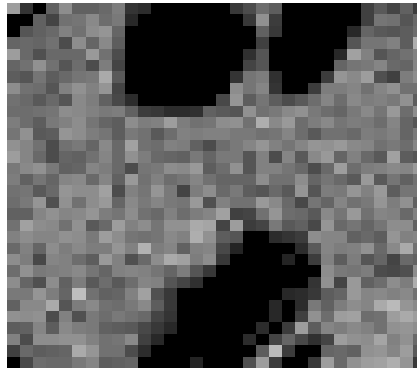
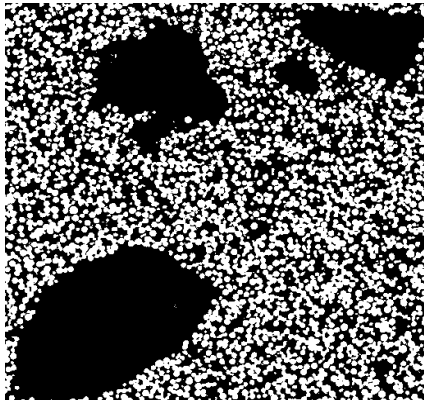


(b) Low resolution observations



(c) Reconstruction, using different types of observations[1]

Real Data



(a) High resolution data used for learning the parameters of the prior model

(b) Low resolution observations

(c) Reconstructed from the posterior the prior model [1]

Summary and Conclusion

- In many pattern recognition tasks, a single observation from one sensor is not sufficient to discover the hidden pattern.
- Data fusion is a key solution for such problems.
- Based on the type of fusion task, i.e. fusion across sensor, domain, attribute and time, variety of data fusion methods can be applied successfully.
- Bayesian approach is a suitable framework in order to model and solve a data fusion problem since
 - It can model the uncertainty observed in the data, and
 - It incorporates the prior belief about the unknown phenomenon under study.

References

- [1] Mohebi, A., Fieguth, P., and Ioannidis, M.A. "Statistical fusion of two-scale images of porous media." *Advances in Water Resources* 32.11 (2009): 1567-1579.
- [2] Mitchell, H. B. *Introduction*. Springer Berlin Heidelberg, 2007.
- [3] Mitchell, H. B. *Image fusion: theories, techniques and applications*. Springer, 2010.
- [4] Stathaki, T. *Image fusion: algorithms and applications*. Access Online via Elsevier, 2011.
- [5] Bennett, E. P., Mason, J. L. and McMillan L. "Multispectral bilateral video fusion." *Image Processing, IEEE Transactions on* 16.5 (2007): 1185-1194.
- [6] Park, S. C., Park, M. K. and Kang, M. G. "Super-resolution image reconstruction: a technical overview." *Signal Processing Magazine, IEEE* 20.3 (2003): 21-36.
- [7] Loza, A., Bull, D., Canagarajah, N., and Achim, A. "Non-Gaussian model-based fusion of noisy images in the wavelet domain." *Computer Vision and Image Understanding* 114, no. 1 (2010): 54-65.
- [8] Pajares, G., and Manuel de la Cruz, j. "A wavelet-based image fusion tutorial." *Pattern recognition* 37.9 (2004): 1855-1872.
- [9] Mohebi, A., and Fieguth, P. "Posterior sampling of scientific images." *Image Analysis and Recognition*. Springer Berlin Heidelberg, 2006. 339-350.
- [10] Mohebi, A., and Fieguth, P. "Statistical fusion and sampling of scientific images." *Image Processing, 2008. ICIP 2008. 15th IEEE International Conference on*. IEEE, 2008.

Thank you!