	raw data		representation			embedded
	in-time-based	transform-based	shape-based	model-based	feature-based	iterative-based
raw data	×	×	×	×	×	×
transformation		X				
representation			X	X 	X 	
clustering	×	X	X	X	×	
example			= +	$\downarrow$ $Y_{t}=\alpha+\beta_{1}Y_{t-1}+\beta_{2}Y_{t-2}+\cdots+\epsilon_{t}$	$\downarrow \\ \bar{x} \\ \sigma^2 \\ \hat{\rho}_k \\ \sum  x_{t+1} - x_t $	Start Initial Values  Expectation  Maximization
algorithm	LKMA	DTW	DFT *	VAR	$eta_{\iota}^{eta_{\iota}}$ : $tsfresh$	No Converged Yes End  GBTM
notes	<ul> <li>low interpretability</li> <li>same interval (if not transformed)</li> <li>same length (no missing)</li> <li>no parameter dependence (ignore temporal order)</li> <li>sensitive to offset (if not transformed)</li> <li>sensitive to noise</li> <li>no shape assumed (incl. sudden changes)</li> <li>fast modeling</li> <li>readily available software</li> <li>cluster interpretation established in the field</li> </ul>		<ul> <li>dangerous if model fit is poor</li> <li>assumptions violated (poor fit)</li> <li>too few observations (poor fit)</li> <li>reduced dimensional space         <ul> <li>use domain knowledge to choose summarize features)</li> </ul> </li> <li>more acurate then raw</li> <li>fast modeling</li> <li>readily available software</li> <li>robust to missing data (b/c calculated on multiple observations)</li> <li>varying intervals</li> <li>varying lengths</li> <li>often scaleable performance (e.g., model only needs to be fitted once)</li> <li>relatively few observations per trajectory</li> </ul>			often assume the same parametric distribution dangerous if model fit is poor (e.g., over- or under fitting) assumptions violated (fit) too few observations (fit) slower with mode complex models robust to missing data (b/c calculated on multiple observations) varying intervals varying lengths interpretatble (e.g., distinct cluster trajectories) good for prediction