

**Aufgabe 2** (Bestimmung einer optimalen Farbpalette mit dem  $k$ -Means-Algorithmus)

- a) Implementieren Sie in einer Programmiersprache Ihrer Wahl den  $k$ -Means-Algorithmus für eine Liste von Eingabevektoren variabler Dimension  $N$ . Die Komponenten der Vektoren seien dabei Gleitkommazahlen. Als Abstandsmaß kommt die Euklidische Distanz zum Einsatz. Initialisieren Sie die  $k$  Clusterzentren zu Beginn der Iteration mit  $k$  zufällig ausgewählten Eingabevektoren. Geben Sie nach jeder Iteration den mittleren quadratischen Abstand der Eingabevektoren zu ihrem aktuell jeweils zugeordneten Klassenzentrum aus. Beenden Sie die Iteration, wenn der Betrag der Differenz des so ermittelten Abstands zu dem der Vorgängeriteration einen einstellbaren Schwellwert unterschreitet, oder alternativ nach einer festen Zahl von Iterationen. Ergebnis ist die Liste der  $k$  Klassenzentren (Codebuch).
- b) Schreiben Sie auf Basis Ihres  $k$ -Means-Algorithmus ein Programm, das durch Clustern der RGB-Werte zu einem jpg-Bild die optimale Farbpalette jeweils mit 4, 16 und 256 Farben bestimmt, und anschließend alle Pixel im Bild durch den zugehörigen Wert aus der ermittelten Farbpalette ersetzt. Die so erzeugten Bilder werden unter dem Dateinamen der Originaldatei, ergänzt um einen entsprechenden Suffix, gespeichert. Aus der Datei `Rosen.jpg` sollen z. B. die Dateien `Rosen-4.jpg`, `Rosen-16.jpg` und `Rosen-256.jpg` erzeugt werden. Verwenden Sie beispielsweise die Testbilder unter `Students/Mustererkennung/Medienverarbeitung-SS2012/Uebung-2/`, oder auch jedes andere Farbbild im jpg-Format.

**Hinweis:** Sie werden feststellen, dass die Berechnung bei hochauflösenden Bildern extrem lange dauert. Die einfachste Lösung hierfür ist es, nicht mit allen Pixeln aus dem Originalbild zu arbeiten, sondern nur mit einem kleinen Bruchteil der Pixel. Sie können z.B. einfach nur jeden hundertsten Pixel als Eingabe für den  $k$ -Means-Algorithmus nutzen, ohne dass sich das Ergebnis nennenswert verändert. Versuchen Sie, einen vernünftigen Kompromiss zu finden!