# 4
# NELDER – MEAD ALGORITHM

This chapter explains the Nelder-Mead optimization algorithm which was responsible for iteratively provides a set of parameter values for each segmentation algorithm. Its objective function was defined by each one of the selected metrics for this study.

## 4.1.
## The Method

Nelder and Mead (Nelder & Mead, 1965) (NM) proposed a stochastic method for the minimization of a function of $n$ variables, which depends on the comparison of function values at $(n+1)$ vertices of a general simplex, followed by the replacement of the vertex with the highest value by another point. The simplex adapts itself to the local landscape, and contracts on to the final minimum. The method is shown to be effective and computationally compact.

Let's $P_0, P_1, \dots, P_n$ be the $(n+1)$ points in $n$-dimensional space defining a simplex, $y_i$ will denote the value of the objective function at $P_i$, and we define:

$$y_h = \max(y_i) \tag{20}$$

$$y_l = \min(y_i) \tag{21}$$

Thus, $P_h$ and $P_l$ correspond to the points where $y_h$ and $y_l$ occur. Further, the centroid of the points with $i \neq h$ is defined as $P_m$. Once $P_l$, $P_h$ and $P_m$ are calculated, an iterative process begins. At each stage in the process, $P_h$ is replaced by a new point. Three operations are executed during the whole process – *reflection*, *contraction* and *expansion*. These operations and each step of the NM algorithm are illustrated in the flowchart of the Figure 7. The process generates a sequence of triangles (which might take different shapes), for which the function values at the vertices get smaller and smaller. The area of the triangles is reduced at each new iteration toward a single point that characterizes the minimum (Mathews & Fink, 2004).

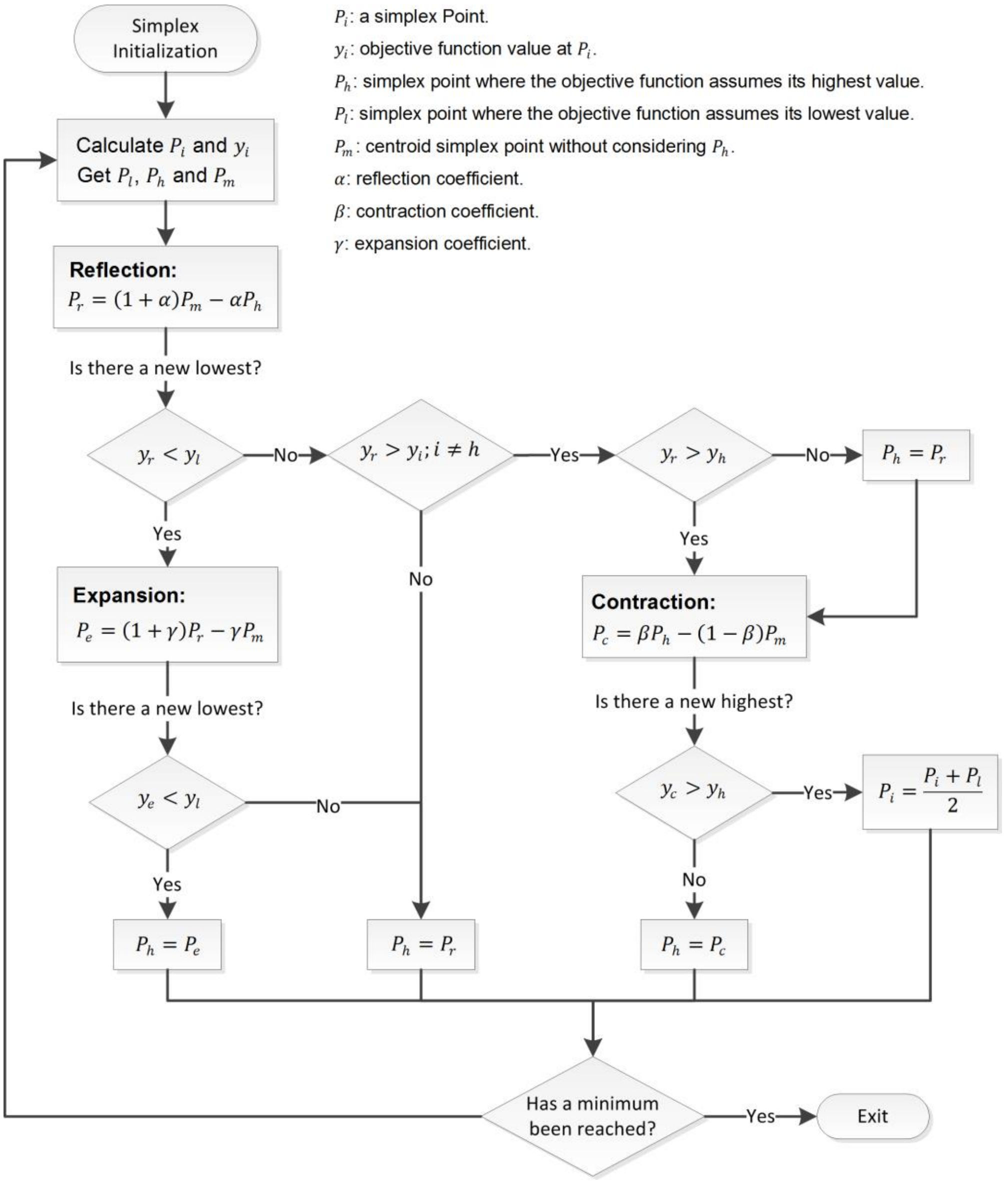


Figure 7: Description of the Nelder – Mead algorithm.

Further details of each step shown in Figure 7 are explained in the following section.

## 4.2. Simplex Initialization

Let’s take a bidimensional example for a better understanding of each operation. Let´s assume that three vertices of the simplex are: $P_i = \{P_1, P_2, P_3\}$, and the values of the objective function in each vertex are: $y_i = y(P_i)$. These values have to be sorted to identify the values of $P_l$ and $P_h$. Also, as there are only three vertices, the remainder value will be defined as $P_s$ to be the point where the objective function assumes its second highest value. Later, the following nomenclature is established:

$$P_l = P_1, \quad P_s = P_2, \quad P_h = P_3 \tag{22}$$

As there are only three vertices, the value of $P_m$ will be defined as follows:

$$P_m = \frac{P_l + P_s}{2} \tag{23}$$

Now, the following operations are defined according to the flowchart in Figure 7.



## 4.3. Reflection

As $P_h$ moves toward $P_l$ and $P_s$, $y$ takes smaller values at points lying away from $P_h$ on the opposite side of the line determined by $P_l$ and $P_s$ (see Figure 8). So, a test point $P_r$ is determined by "reflecting" the triangle through the line $\overline{P_l P_s}$ (see Figure 8). Then, the value of $P_r$ is calculated as follows:

$$P_r = (1 + \alpha)P_m - \alpha P_h \tag{24}$$

where $\alpha$ is a positive constant called the *reflection coefficient*.

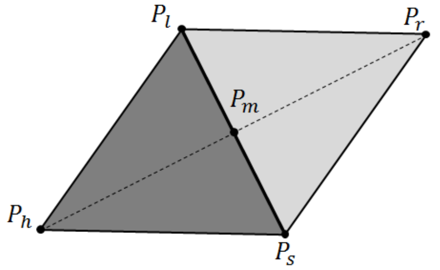


Figure 8: Reflection Operation (modified from (Mathews & Fink, 2004)).

## 4.4. Expansion

If after a reflection operation, $y_r$ is smaller than $y_h$, then, it was a successful reflection toward the minimum. Perhaps the minimum is just a bit farther than the point $P_r$. Thus, the line segment through $P_m$ and $P_r$ is extended to the point $P_e$ (see Figure 9). If $y_e$ is less than $y_r$, it has been a good expansion and $P_h$ is replaced by $P_e$. Then, the value of $P_e$ is calculated as follows:

$$P_e = (1 + \gamma)P_r - \gamma P_m \quad \textbf{(25)}$$

where $\gamma$ is a constant greater than unity and is called the *expansion coefficient*.

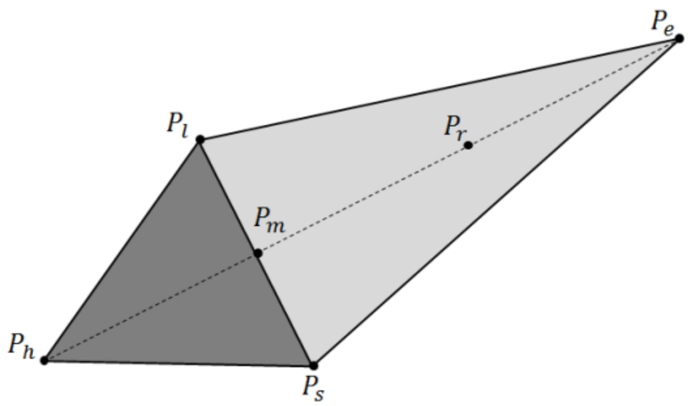


Figure 9: Expansion Operation (modified from (Mathews & Fink, 2004)).



## 4.5. Contraction

If $y_r$ is equals to $y_h$, it is necessary to test another point. Perhaps the objective function is smaller at $P_m$. However, the replacement of $P_r$ by $P_m$ would imply transforming the triangle $\overline{P_r P_l P_s}$ into a line, which is not desirable. Instead, we take the midpoints $P_{c1}$ and $P_{c2}$ of the line segments $\overline{P_h P_m}$ and $\overline{P_m P_r}$ respectively (see Figure 10). These points are calculated as follows:

$$P_c = \beta P_h - (1 - \beta)P_m \quad \textbf{(26)}$$

where $\beta$ lies between 0 and 1 and is called the *contraction coefficient*.

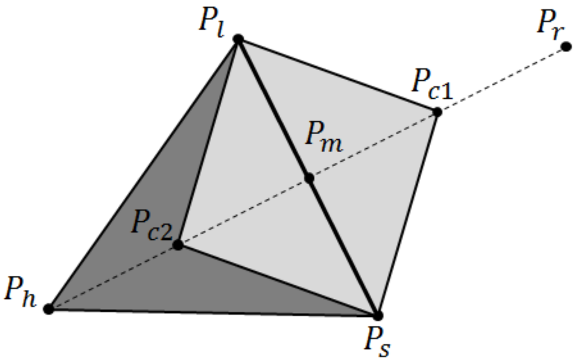


Figure 10: Contraction Operation (modified from (Mathews & Fink, 2004)).

If the function value at $P_c$ is not less than the value at $P_h$, a failed contraction has occurred. In that case, the points $P_h$ and $P_s$ must shrink toward $P_l$. The point $P_s$ is then replaced by $P_m$ and the point $P_h$ by $P_{m1}$, which is the midpoint of the line segment $\overline{P_h P_l}$ (see Figure 11).

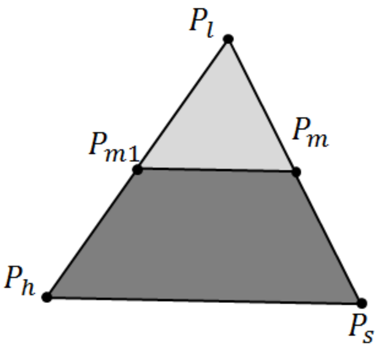


Figure 11: In case of a failed contraction, shrinking the triangle towards $P_l$ is done (modified from (Mathews & Fink, 2004)).

These operations aim to improve the computational efficiency. This is achieved by reducing the number of evaluations of the objective function because it is only evaluated at the beginning of each operation. At each step, the algorithm searches for a point that minimizes the objective function by evaluating this function at the vertices of the simplex that is generated at each iteration. When $P_h$ is found, the algorithm terminates the current step and updates the vertices of the simplex. This procedure is repeated iteratively to find the optimal solution or until the maximum number of iterations is exceeded.